DOCUMENT RESUME

ED 415 691 FL 024 999

AUTHOR	Griffee, Dale T., Ed.; Nunan, David, Ed.
TITLE	Classroom Teachers and Classroom Research. JALT Applied Materials.
INSTITUTION	Japan Association for Language Teaching, Tokyo.
PUB DATE	1997-00-00
NOTE	151p.; For individual articles, see FL 025 000-011.
PUB TYPE	Collected Works - General (020)
EDRS PRICE	MF01/PC07 Plus Postage.
DESCRIPTORS	Annotated Bibliographies; *Classroom Research; *Classroom Techniques; Ethnography; Foreign Countries; Higher Education; Interviews; Literature Reviews; Proposal Writing; Qualitative Research; Questioning Techniques; Research Design; *Research Methodology; Research Proposals; Second Language Instruction; Second Language Programs; *Second Languages; Statistical Analysis; Students; Surveys; Teacher Role

ABSTRACT

This collection of papers leads classroom language teachers through the process of developing and completing a classroom research project. Arranged in four sections, they include: "Language Teaching and Research" (David Nunan); "Where Are We Now? Trends, Teachers, and Classroom Research" (Dale T. Griffee); "First Things First: Writing the Research Proposal" (Martin White, Diane White); "Literature Reviews: Obtaining Perspective" (Paul Riley); "Designing a Language Study" (James Dean Brown); "Methods for the 'Research Challenged'" (Robert Homan, Chris Poel); "Action Research: Something for Everyone" (Greg Hadley); "Students as Ethnographers" (Linda Donan); "Designing Surveys for Language Programs" (Brown); "Bare Minimum Knowledge for Understanding Statistical Research Studies" (Nagata Hiroto); "In-Depth Interviewing as Qualitative Investigation" (Marilyn Books); and "Teachers and Research: An Annotated Bibliography" (Griffee). MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 415 691

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

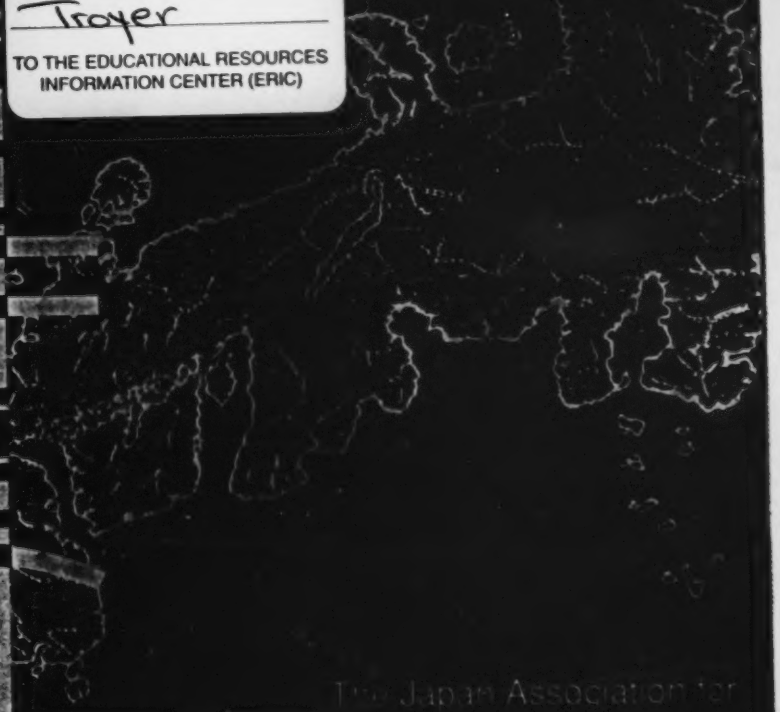
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Gene Van Troyer

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

JALT APPLIED MATERIALS



BEST COPY AVAILABLE

2

The Japan Association for Language Teaching

The JALT Applied Materials Series

Series editor, Dale T. Griffee

This new series offers collections of papers of interest to classroom teachers. JAM books combine current research, theory, and classroom application of particular interest in Japan and other countries in Asia.

In this series:

Language Testing in Japan, edited by James Dean Brown and Sayako Yamashita.

Classroom Teachers and Classroom Research, edited by Dale T. Griffee and David Nunan.

Scheduled for release:

Cooperative Learning, edited by David Kluge and Steve McGuire.

Second Language Acquisition, edited by Peter Robinson, Steven Ross, and Mark Sawyer.

JALT Applied Materials

Classroom Teachers and Classroom Research

Dale T. Griffee & David Nunan

Dale T. Griffee, Series Editor

Lyneve Rappell, Production Manager/Assistant Editor

The Japan Association
for Language Teaching
全国語学教育学会
Tokyo, Japan

4

JALT Applied Materials

A series sponsored by the Publications Board
of the Japan Association for Language Teaching.

JALT President: Gene van Troyer
Publications Board Chair: Steve Cornwell
Cover design and graphics: Junji Maeda, Press Associates, Nagoya
Design and typography: The Word Works, Ltd., Yokohama
Printing and binding: Koshinsha K.K., Osaka

CLASSROOM TEACHERS AND CLASSROOM RESEARCH

copyright © 1997 by the Japan Association for Language Teaching,
Dale T. Griffee and David Nunan.

All rights reserved. Printed in Japan. No part of this book may be used or reproduced in any form whatsoever without written permission of the editors, except in cases of brief quotations embodied in scholarly articles and reviews. For information contact: JALT Central Office, Urban Edge Bldg. 5th Fl. 1-37-9 Taito, Taito-ku, Tokyo 110, Japan.

Cataloging-in-Publication Data

Griffee, Dale T., & Nunan, David (Eds.)
Classroom Teachers and Classroom Research.

Bibliography: p.
1. Applied linguistics—Classroom Research. I. Title
1997
ISBN 4-9900370-2-6

Contents

Forward 7
About the Contributors 9

SECTION I: AN OVERVIEW

1. Language Teaching and Research 13
David Nunan
2. Where Are We Now? Trends, Teachers, and Classroom Research 23
Dale T. Griffee

SECTION II: GETTING STARTED

3. First Things First: Writing the Research Proposal 39
Martin White and Diane White
4. Literature Reviews: Obtaining Perspective 49
Paul Riley
5. Designing a Language Study 55
James Dean Brown
6. Methods for the Research Challenged 71
Robert Homan and Chris Poel

SECTION III: RESEARCH METHODS

7. Action Research: Something for Everyone 87
Greg Hadley
8. Students as Ethnographers 99
Linda Donan
9. Designing Surveys for Language Programs 109
James Dean Brown
10. Bare Minimum Knowledge
for Understanding Statistical Research Studies 123
Nagata Hiroto
11. In-depth Interviewing as Qualitative Investigation 137
Marilyn Books

SECTION IV: FUTURE READING

12. Teachers and Research: An Annotated Bibliography 149
Dale T. Griffee

Forward

About JALT Applied Materials

Welcome to the JALT Applied Materials (JAM) series, of which *Classroom Teachers and Classroom Research* is the second volume. The concept behind JAM is that each proposed volume is a collection of papers focused on a single theme, written by classroom teachers, some experienced and some just beginning.

The Situation

Why begin a new series? At present, books and collections of papers on various subjects are written and published by authors from various parts of the world, mainly from English-speaking countries such as Australia, Canada, New Zealand, the U.K. and the U.S. The authors of these books and articles are typically university professors who teach courses in graduate programs and speak at conferences around the world including JALT's annual international conference. There exists, however, a gap between the international level where these books are generated and written and the local teaching community in Japan and Asia. While these books are of general professional interest, local teachers are usually not sought as contributors, and only find out about the book when they see it in a book store or catalog. The JAM series was created too help fill this gap.

Purpose

The JAM series is targeted at: (a) improving the quality of research and academic writing in Japan and Asia; (b) publishing collections of theoretically grounded, reader-friendly, and classroom oriented articles on subjects of interest to classroom teachers in Japan and Asia; (c) giving Japan and Asia based classroom teachers a publication outlet not heretofore available; and (d) helping teachers around the world implement new ideas in their classrooms.

How JAM Works

We feel that JALT has reached a level of development where many of its members are capable of writing creative and professional papers. Research and writing, however, are not easy tasks, and we feel that JALT members would be helped to produce quality papers if they had comprehensive and constructive editorial support and guidance. To this end, the JAM series is guided by an editorial network composed of a general series editor, one or more local editors, and the possibility of an international editor for each collection of papers, all of whom would work with the writers of that collection.

Any person who is a resident of Japan, a member of JALT, and primarily a classroom teacher may request to be a JAM editor. Persons requesting to be editors would be asked to submit their CV and publications relevant to their subject. This Japan-based local editor (or editors) could, if they desired, ask an international expert in the field to assist them. We will refer to this editor as the "global" editor. After approval by the JAM committee of the theme and editors, the next step would be for the local editor to issue a call for abstracts. The date for first drafts would be set by the local editor and the series editor. All manuscripts must be approved by all editors: first by the local editor, then by the global editor, and finally by the series editor. Thus, JAM is a vetted publication, and it will be a personal as well as professional achievement to have an article published in a JAM collection.

Contributing authors to a JAM collection are expected to be classroom teachers in Japan or a country with similar teaching conditions (e.g., Korea, China, Taiwan, or Thailand), to be familiar with JALT and its mission, and to submit an original article which has not been submitted to another publication. JAM welcomes new as well as experienced writers, and JAM authors need not have published previously in the theme area.

Conclusion: A Unique Opportunity

JALT is a professional organization of teachers who occupy the niche between new developments and traditional classroom practices. As such, we are historically positioned to creatively interact between theory and the daily reality of the classroom. We firmly believe that what we have to report will be of significant value to teachers throughout Asia and around the world. We are confident that you will concur.

N.B. Persons who desire further information about JAM should contact the series editor, Dale T. Griffee at Seigakuin University, 1-1 Tosaki, Ageo-shi, Saitama-ken 362, Japan.

About the Contributors

Marilyn Books obtained her Ph.D. degree in 1995 in intercultural communication, specializing in Japanese and American communication styles. She worked toward the doctoral degree and a Master of Arts diploma while in Japan. Now she is tackling Japanese language study. She is a part-time lecturer at Meiji University and Waseda University, among others. A Canadian, she has been in Japan for ten years.

James Dean ("JD") Brown, Professor on the graduate faculty of the Department of ESL at the University of Hawaii at Mania, specializes in language testing, curriculum design, program evaluation, and research methods. He has taught in France, the People's Republic of China, Saudi Arabia, Japan, Brazil, and the United States. He has served on the editorial boards of the *TESOL Quarterly* and *JALT Journal*, as well as on the TOEFL Research Committee, TESOL Advisory Committee on Research, and Executive Board of TESOL. In addition to numerous journal articles and book chapters, he has published three books entitled: *Understanding Research in Second Language Learning: A Teacher's Guide to Statistics and Research Design* (Cambridge, 1988); *The Elements of Language Curriculum: A Systematic Approach to Program Development* (Heinle &Heinle, 1995); and *Testing in Language Programs* (Prentice-Hall, 1995).

Linda Donan

Dale T. Griffee is the series editor for the JALT Applied Linguistics series. He is Associate Professor at Seigakuin University, Saitama-ken. He graduated from Baylor University in Waco, Texas; received his M.A. in TESOL from the School for International Training, Brattleboro, Vermont; and is enrolled in the Temple University Japan doctoral program. His research interests are curriculum, testing, and evaluation in the learner centered classroom.

Greg S. Hadley teaches English at Keiwa College, a liberal arts Christian College in Niigata prefecture. His current interests include Sociolinguistics and Psycholinguistic

BEST COPY AVAILABLE

research involving Personal Construct Repertory Grids. He is finishing his M.A. in TESL/TEFL through the University of Birmingham, England.

Robert M. Homan is a member of the English department at International Christian University. His current interests are in developing cooperative learning tasks, and writing textbooks.

Nagata Hiroto worked for more than a decade as a language program coordinator at such Japanese companies as Matsushita Electric Industrial, Marubeni, and Nissan. He has published a number of articles on language teaching, and a variety of books and dictionaries for both children and adult learners of English. He received his M.Ed. in TESOL at Temple University Japan, where he also finished his doctorate. He currently conducts classes at Yokohama National, Aoyama Gakuin, and Soka Universities.

David Nunan is director of the English Centre and professor of English at the University of Hong Kong. His areas of research include teacher training and classroom research. His articles have appeared in major journals around the world, and he is a frequent speaker at conferences. His books include *Research Methods in Language Learning* and *The Learner-Centred Classroom* and most recently *Voices from the Language Classroom*, a collection of articles edited with Kathleen Bailey.

Chris Poel is an associate professor of English at Musashi Institute of Technology, Yokohama. Since he started teaching in Japan in 1980, he has developed interests in task-based language teaching, cooperative learning, university entrance exams, and materials design.

Paul Riley is an editor in the East Asia Publishing Group of Oxford University Press (OUP). Before joining the editorial staff, he was an ELT Consultant in the OUP Tokyo office, Manager of International Training and Education at Isuzu Motors, and a teacher and teacher trainer at various schools in the Kanto region and Nagoya. He received an M.Ed. in TESOL from Temple University Japan.

Diane White has been teaching foreign languages since 1976 in New Zealand, France, Okinawa, and Tokyo. Now at Tokyo Woman's Christian University, she teaches courses on spoken and written English, including one on Research Writing. Her own research interests lie in the field of language acquisition, and most particularly in genre analysis and translation. She is currently pursuing another M.A. degree, this time in Japanese.

Martin White received his M.Ed. in TESOL from Temple University Japan in Tokyo. He has taught in Britain, Germany, Malaysia, Japan, and Saudi Arabia. He is currently working in the International Communications Program of AIG KK. He is a member of JALT and has held several offices in the Tokyo JALT chapter. His interest areas are reading, discourse analysis and materials development.

SECTION I

An Overview

Chapter 1

Language Teaching and Research

David Nunan

University of Hong Kong

In this introductory paper, I should like to do three things. In the first place, I should like to articulate a vision of research which is in harmony with second and foreign language education. Secondly, I should like to summarize what I see as some of the things we have learned from research that can help us as we struggle to improve what we do in the classroom. The final thing I should like to do is to set out what I see as some important future trends for applied research in language education.

Approaches to Research

It is something of a curiosity to me that few of the recent spate of books on research in language education attempt to offer a definition of the term itself. So, years ago I asked a group of my graduate students, all language teachers who were embarking on an M.A. in Applied Linguistics, what the term meant for them. Some of their responses are set out below.

- Enquiry has two components: process and product. The process is about an area of enquiry and the process used to pursue that. The product is the knowledge generated from the process as well as the initial area to be presented.
- A process which involves (a) defining a problem, (b) stating an objective, and (c) formulating an hypothesis. It involves gathering information, classification, analysis, and interpretation to see to what extent the initial objective has been achieved.

- Undertaking structured investigation which hopefully results in greater understanding of the chosen interest area. Ultimately, this investigation becomes accessible to the public.
- An activity which analyses and critically evaluates some problem.
- To collect and analyse the data in a specific field with the purpose of proving your theory.
- Evaluation, asking questions, investigations, analysis, confirming hypotheses, overview, gathering and analysing data in a specific field according to certain predetermined methods.

For me, research is a systematic process of inquiry in which the researcher poses a question or questions, collects relevant data on the question(s), analyses and interprets the questions, and makes the results of the inquiry publicly accessible in some way.

Donald Freeman, who has written extensively on research in TESOL is one person who has offered a characterization of research suggesting that it is "a basic process of developing and rendering viable interpretations for things in the world" (Freeman, 1996, p. 102). While this is a normal and natural part of everyday life, research differs from normal human curiosity through its systematicity and accessibility to scrutiny. In practical terms, research is a process of formulating questions or articulating "puzzles" relating to practice, collecting relevant data that might have a bearing on such questions or puzzles, interpreting and explaining the data, and making the results of the inquiry publicly accessible in some way (Nunan, 1992a). As Freeman pointed out, to give this process the label "research" is important because it gives the activity value: "When [the] questioning of practice takes place within a framework labeled 'research,' understanding the complexity of teaching can become a public and legitimate part of being a teacher" (Freeman, 1996, p. 103).

Many commentators argue that there are two competing traditions in research. The first of these, the quantitative tradition, is obtrusive and controlled and is concerned with established generalizable relationships between variables. The other, the qualitative tradition, is concerned with generating insight and understanding rather than establishing "truths." More recently, it has been argued that this distinction is simplistic and naive. Elsewhere, I have suggested that while the distinction was simplistic it is also persistent:

> One reason for the persistence of the distinction between quantitative and qualitative research, is that the two approaches represent different ways of thinking about and understanding the world around us. Underlying the development of different research traditions and methods is a debate on the nature of knowledge, and the status of assertions about the world, and the debate itself is ultimately a philosophical one. It is commonly assumed that the function of research is to add to our knowledge of the world and to demonstrate the 'truth' of the commonsense notions we have about the world. . . . In developing one's own philosophy on

> research, it is important for us to determine where we stand on the notion of 'truth' in relation to research. What is truth? (Even more basically, do we accept that there is such a thing as 'truth'?) What is evidence? Can we ever 'prove' anything? What evidence would compel us to accept the truth of an assertion or proposition? These are questions which need to be borne constantly in mind as one reads and evaluates research. (Nunan, 1992a, p. 10)

What Have We Learned From Research?

In conducting teacher education seminars on research methods, I am sometimes confronted with the view that research is an esoteric activity that has little to do with the real world, and little to say to the classroom practitioner. I beg to differ. While honest research is messy, the results often inconsistent if not contradictory, and the outcomes uncertain, research has helped to advance our understanding of the processes underlying language acquisition. I also believe that research can offer us an empirical basis for practice. At a recent conference in Brazil (Nunan, 1996b), I was asked to address the relationship of research to practice, and to spell out the practical implications of research for language pedagogy. I gave my own idiosyncratic list of principles that guide my teaching and materials writing that are derived from my own research and that of others. These I have reproduced below, along with the research from which they are derived.

1. Make instructional goals explicit to learners in ways they can understand.

- Goal setting can have exceptional importance in stimulating L2 learning and enhancing motivation (Oxford & Shearin, 1993).
- Motivation is enhanced when learning goals are made clear (Reilly, 1994).
- Learning was enhanced in classes where teachers made goals explicit. Explicit goal setting was relatively rare in classrooms observed (Nunan, 1996a).

2. Give learners opportunities to make their own contributions to the learning process.

- Learners have definite views on what they want to learn and how they want to learn. These are often at variance with the views of the teacher. Learning is enhanced when learners are involved in making choices (Nunan, 1987).
- Young learners are able to take responsibility for planning, organising, managing and evaluating their own learning (Dam & Gabrielsen, 1988).
- Students want to be involved in the selection of language content and the learning process. There are major mismatches between the desires of the students and the mandated curriculum (Widdows & Voller, 1991).
- Classroom topics nominated by learners were much more likely to be learned than topics determined by the teacher (Slimani, 1992).
- Learning is enhanced when students are actively involved in selecting content, learning tasks, and evaluation (Heath, 1992).

3. Encourage active communication through sequenced, achievable tasks.

- Spoken output in groupwork is a significant factor to the acquisition of vocabulary (Hall, 1991).
- Learner participation in class relates significantly to improvements in language proficiency (Lim, 1992).
- Motivation is enhanced when tasks are sequenced and linked in ways that make sense to the learner (Reilly, 1994).
- Active use of target language with strong emphasis on practice in naturalistic situations is notable in helping students attain higher proficiency (Green & Oxford, 1995).

4. Provide opportunities for learners to apply their skills beyond the classroom.

- Supplementing classroom instruction with out-of-class tasks results in significantly increased language gains (Montgomery & Eisenstein, 1985).
- Instruction and opportunities to communicate out of class are both necessary. Improvement occurred when subject consciously "noticed the gap" (Schmidt & Frota, 1986).
- The "good" foreign language learner finds ways of activating his or her language out of class (Nunan, 1991).

5. Teach learning strategies as well as language content.

- The ability to infer or induct rules is an important aspect of language aptitude (Carroll, 1981).
- Strategy training had a significant effect on speaking development (O'Malley, Chamot, Stewner-Manzanares, Russo, & Kupper, 1985).
- Effective learners are aware of the processes underlying their own learning and seek to use appropriate learning strategies to control their own learning (Jones, Palincsar, Ogle, & Carr, 1987).
- More effective learners use strategies more frequently, and use a greater variety of strategies than students designated as less effective learners (O'Malley & Chamot, 1990).
- Good learners are aware of strategies that worked for them, and can articulate these (Nunan, 1991).
- There is a significant relationship between strategy use and success in language learning (Green & Oxford, 1995).

6. Teach grammar in ways that show the essential harmony between form and function.

- Grammar exists in order to enable learners to communicate in increasingly sophisticated ways (Nunan, 1993).
- Grammar and vocabulary are best acquired through learner involvement in the processing and production of discourse (McCarthy & Carter, 1994).

7. Go beyond declarative knowledge to procedural skills development.

- "Communicative" classrooms with instruction plus opportunities for interaction are superior to "traditional" instruction and also to immersion programs (Spada, 1990).
- Formal instruction resulted in acquisition of some structures (passives) but not others (tense and aspect). Explicit (declarative) knowledge can be converted to implicit (procedural) knowledge through practice (Zhou, 1991).
- Declarative knowledge (ability to identify errors and state rule violation) does not lead to procedural knowledge (ability to put known forms to communicative effect) without opportunities to activate knowledge through output activities (Wudong, 1994).

8. Give learners an opportunity to work with authentic data.

- Non-authentic data misrepresents the nature of genuine communication (Nunan, 1991).
- Authenticity significantly enhances motivation in foreign language classes (Ho, 1995).

9. Maximize opportunities for learners to work cooperatively.

- Small group tasks prompt students to use a greater range of language functions than teacher fronted tasks (Long, Adams, & Castanos, 1976).
- Cooperative learning leads to positive interdependence but also individual accountability, extensive face-to-face interaction, and the development of social skills (Kohonen, 1992).

10. Provide learners with opportunities to self-monitor and self-check.

- Opportunities for learners to self-monitor and self-check lead to greater sensitivity to the learning process. Learners develop skills in articulating what they want to learn and how they want to learn (Nunan, 1995).

Future Trends

And so to the future. What direction would I like to see research in language education taking? I believe that a future research agenda should be contextualised and classroom oriented. It should forge closer links between teaching and research of the type indicated in the body of this paper. It should encourage collaboration between researchers, teachers and students. And, finally, it should seek to extend our vision of research. Each of these points is commented on below.

1. Contextualized and classroom-oriented.

The need for research to be contextualized is slowly being recognised. Even university-based researchers with little direct contact to classroom realities are beginning to recognise the limited applicability of research outcomes which

have been derived from contexts other than those in which teaching and learning typically occur. To acknowledge that context affects behavior is to acknowledge that conclusions reached in one context can be taken as nothing more than working hypotheses in other contexts.

2. Closer links between teaching and research.
Related to the first point is an appreciation of the need for developing closer links between teaching and research. Allwright and Bailey (1991) coined the phrase "exploratory teaching" to capture this alternative vision in which it is unnecessary to decide whether one is engaged in teaching or research; rather one is engaged in both. In other words, exploratory teaching describes a philosophical stance or attitude of mind towards one's classroom practice. It is an ongoing process of working constantly for deeper understanding and increasing effectiveness in the classroom. Underlying their approach to reflective teaching are the assumptions that (a) autonomous teachers are the key link between teaching, learning, and research, and (b) becoming more effective as a teacher is a life-long, spiraling process.

> The teacher is the researcher's link with learners, and also the learners' link with research. The teacher is contracted to help learners learn, but can do so better by knowing about previous research and by using the procedures of classroom research to understand better what is happening in his or her own classroom. In this way the exploratory teacher will not only improve achievement but will also contribute to our general research knowledge about how language classrooms work. This is what we mean by 'exploratory teaching'—teaching that not only tries out new ideas, but that also tries to learn as much as possible from doing so. . . . Any good experienced teacher will no doubt spend a lot of class time on ideas that are tried and trusted. Turning that 'good' teaching into 'exploratory' teaching is a matter of trying to find out what makes the tried and trusted ideas successful. (Allwright & Bailey, 1991, p. 197)

3. Collaborative research
Collaborative research is research in which all those involved in the research process, teacher-researchers, administrators, and informants have an active role and a voice within the process itself. The term also refers to inquiry carried out by multiple participants across a number of teaching sites. The value of collaboration is slowly being recognised, and is something which the field should seek to encourage. Such collaboration can be facilitated by the networks and systems (electronic and otherwise), which can help to link together individuals who might be working in different contexts (and even different countries). Teacher-researchers who collaborate can generate more powerful ideas, initiate more interesting research, and achieve more useful outcomes, than those working in isolation. A useful question for consideration is: "What are the

central characteristics of a collaborative approach to classroom research, and in what ways do the different contributions of teachers, learners and researchers provide us with insights which would be difficult to obtain in any other way?" (Nunan, 1992b, p. 8).

4. An extended vision of the nature of research.
There is growing acceptance within the social sciences generally that researchers need to develop a greater range of models, tools, and analytical techniques. Teacher-researchers working in language education should develop an extended vision of the nature of research. They should extend the theoretical bases of the research as well as the range of research tools, techniques, and methods. While conventional research methods—experiments, ethnographies, case studies and the like—have their place within a practitioner-oriented research agenda, there is a need to go beyond these. Ways of extending the agenda exist in the literature. These include focused teaching (Allwright & Bailey, 1991), action research (Nunan, 1989), narrative accounts (Freeman, 1994).

The four key principles set out in this section underlie my vision of the kind of research agenda I would like to develop and promote. Such an agenda, while not rejecting quantitative research, seeks to enhance the status of constructivist, humanistic research paradigms which are focused on understanding and explaining the contexts in which language education is conducted around the world. Insight and understanding, rather than causality and proof capture the essence of the perspective I am trying to promote.

References

Allwright, D., & Bailey, K. (1991). *Focus on the language classroom*. Cambridge: Cambridge University Press.

Carroll, J. B. (1981). Twenty-five years of research on foreign language aptitude. In K. C. Diller (Ed.), *Individual differences and universals in language learning aptitude* (pp. 83–118). Rowley, MA: Newbury House.

Dam, L., & Gabrielsen, G. (1988). Developing learner autonomy in a school context: A six-year experiment beginning in the learners' first year of English. In H. Holec (Ed.), *Autonomy and self-directed learning: Present fields of application*. Strasbourg: Council of Europe.

Freeman, D. (1994, October). The storyteller, the anthropologist, the theoretician, and the social activist: Charting the dimensions of teacher-research. Plenary presentation at TESOL France.

Freeman, D. (1996). Redefining the relationship between research and what teachers know. In K. Bailey and D. Nunan (Eds.), *Voices from the language classroom*. Cambridge: Cambridge University Press.

Green, J. M., & Oxford, R. (1995). A closer look at learning strategies, L2 proficiency, and gender. *TESOL Quarterly, 29*(2), 261–297.

Hall, S. (1991). *The effect of split information tasks on the acquisition of mathematical*

vocabulary. Unpublished masters dissertation, Victoria University of Wellington, Wellington, New Zealand.

Heath, S. B. (1992). Literacy skills or literate skills? Considerations for ESL/EFL learners. In D. Nunan (Ed.), *Collaborative language learning and teaching* (pp. 40–55). Cambridge: Cambridge University Press.

Ho, A. (1995). *Task motivation in language learning*. Unpublished master's thesis, Chinese University of Hong Kong, Hong Kong.

Jones, B., Palinscar, A., Ogle, D., & Carr, E. (1987). *Strategic teaching and learning: Cognitive instruction in the content areas*. Alexandria, VA: Association for Supervision and Curriculum Development.

Kohonen, V. (1992). Experiential language learning: Second language learning as cooperative learner education. In D. Nunan (Ed.), *Collaborative language learning and teaching* (pp. 14–39). Cambridge: Cambridge University Press.

Lim, S. (1992). *Investigating learner participation in teacher-led classroom discussions in junior colleges in Singapore from a second language acquisition perspective*. Unpublished doctoral dissertation, National University of Singapore, Singapore.

Long, M. H., Adams, L., & Castanos, F. (1976). Doing things with words: Verbal interaction in lockstep and small group classroom situations. In R. Crymes, & J. Fanselow (Eds.), *On TESOL '76*. Washington, DC: TESOL.

McCarthy, M., & Carter, R. (1994, March). *The grammar of spoken discourse*. Paper presented at the Annual TESOL Convention, Baltimore.

Montgomery, C., & Eisenstein, M. (1985). Real reality revisited: An experimental communicative course in ESL. *TESOL Quarterly, 19* (2), 317–335.

Nunan, D. (1987, April). *Hidden agendas in the language classroom*. Paper presented at the RELC Regional Seminar, Singapore.

Nunan, D. (1989). *Understanding language classrooms*. London: Prentice-Hall.

Nunan, D. (1991). *Language teaching methodology*. London: Prentice-Hall.

Nunan, D. (1992a). *Research methods in language learning*. Cambridge: Cambridge University Press.

Nunan, D. (Ed.) (1992b). *Collaborative language learning and teaching*. Cambridge: Cambridge University Press.

Nunan, D. (1993). *Introducing discourse analysis*. London: Penguin.

Nunan, D. (1995a). *ATLAS: learning-centered communication*. Boston: Heinle & Heinle.

Nunan, D. (1995b). Self-assessment as a tool for learning. In D. Nunan, R. Berry, & V. Berry (Eds.), *Bringing about change in language education* (pp. 41–44). Hong Kong: University of Hong Kong.

Nunan, D. (1995c). Hidden voices: Insiders' perspectives on classroom interaction. In K. Bailey, & D. Nunan (Eds.), *Voices from the language classroom: Qualitative research in second language education* (pp. 41–56). New York: Cambridge University Press.

Nunan, D. (1998b, July). *Ten ideas that changed the way I teach*. Plenary presentation, BrazTESOL Convention, Goiania, Brazil.

O'Malley, J. M., Chamot, A. U., Stewner-Manzanares, G., Russo, G., & Kupper, L. (1985). Learning strategy applications with students of English as a second language. *TESOL Quarterly, 19*(2), 285–286.

O'Malley, J. M., & Chamot, A. U. (1990). *Learning strategies in second language acquisition*. Cambridge: Cambridge University Press.

Oxford, R., & Shearin, J. (1993). Language learning motivation: Expanding the theoretical framework. *Modern Language Journal, 78*, 12–28.

Reilly, P. (1994). *The effect of teacher strategies on students' motivation levels in English language classes*. Unpublished master's thesis, University of the Americas, Mexico City, Mexico.

Schmidt, R., & Frota, S. (1986). Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese. In R. Day (Ed.), *Talking to learn: Conversation in second language acquisition* (pp. 237–326). Rowley, MA: Newbury House.

Slimani, A. (1992). Evaluation of classroom interaction. In J. C. Alderson, & A. Beretta (Eds.), *Evaluating second language education* (pp. 197–211). Cambridge: Cambridge University Press.

Spada, N. (1990). Observing classroom behaviors and learning outcomes in different second language programs. In J. C. Richards, & D. Nunan (Eds.), *Second language teaching education* (pp. 293–310). Cambridge: Cambridge University Press.

Widdows, S., & Voller, P. (1991). PANSI: A survey of ELT needs of Japanese university students. *Cross Currents XVIII*, 2.

Wudong, W. (1994). English language development in China. Unpublished doctoral dissertation, University of Tasmania, Hobart, Australia.

Zhou, Y. P. (1991). The effect of explicit instruction on the acquisition of English grammatical structures by Chinese learners. In C. James, & P. Garrett (Eds.), *Language awareness in the classroom* (pp. 254–277). London: Longman.

BEST COPY AVAILABLE

Chapter 2

Where Are We Now? Trends, Teachers, and Classroom Research

Dale T. Griffee

Seigakuin University

Imagine that you have been hired by JR (Japan Rail) to teach a special class on one of their newest express trains. You teach in a special classroom which has been built in the train. You glance out the window as the train pulls out of the station and then turn your attention to your class. The class goes well and the hours fly by. You take a break and look out the window, but the scenery is different and you wonder to yourself, where are we now?

As the Japan Association for Language Teaching (JALT) faces the 21st century and the beginning of the new *JALT Applied Materials* (JAM) series, it seems appropriate to ask the question, where are we now and by implication, where are we going? The working hypothesis for this article is that being a language teacher in Japan is like riding a fast moving train. Every so often it is a good idea to look out the window and reflect on just where we are, where we are going, and perhaps what we might do when we arrive. This article will deal with three questions. The first question is, what trends are being noticed and articulated by the teaching profession at large?; the second question is, to what extent does the JALT leadership agree or disagree with these trends?; and the third question is, what additional trends do JALT leaders perceive?

Definitions

By the word "trend," I mean an issue that is identified as a trend by one or more sources in the literature. By the term "teaching profession at large," I mean those persons teaching English as a second language anywhere in the world including Japan, especially those persons who are members of teaching organizations such as TESOL, IATEFL, and JALT, and who write articles that appear in the journals such as those published by these organizations. By the term "we" as in the sentence, "Where are we now?" I include the teaching profession at large with special emphasis on JALT members. For the purpose of this article, the term "JALT leadership," will be taken to mean those persons listed as chapter officers, SIG officers, elected national officers, and key committee persons in the April, 1994, supplement to The Language Teacher. This is not an altogether inclusive definition of JALT leadership but it has the advantage of being a relatively small group, it is a group whose names and addresses are publicly listed, and because of the offices these individuals hold they can reasonably be expected to be interested in the issue of professional trends.

The Trends

My first question is, what trends are being noticed and attended to by the teaching profession at large? From the literature, 15 trends were identified.

1. As a knowledge-based industry, we are in the process of becoming part of the global information network (Ashworth, 1991; Bowers, 1994; Fanselow, 1987; Schinke-Llano, 1991; Widdowson, 1986, 1992).
2. We are experiencing dramatic growth in numbers in our field (Alatis, 1987; Ashworth,1991; Schinke-Llano, 1991; Swales, 1993).
3. The new *Monbusho* (Ministry of Education, Science and Welfare) course guidelines will have a significant impact (Aiga, 1990; Carter, Goold, & Madeley, 1993; Goold, Madeley, & Carter, 1993; Wada, 1994, 1995; Wada & Cominos, 1994; Wadden, 1994).
4. The emergence of U. S. colleges and universities in Japan (Helgesen, 1991).
5. The JET, AET program will have an influence (Aiga, 1990; Cominos, 1992; Leedham, 1991; Shang-Ikeda, 1992; Wada, 1994,1995; Wada & Cominos, 1994; Yukawa, 1992).
6. More teachers are becoming interested in classroom based research (Allwright, 1994; Gaies, 1991; Genesee, 1994; Grosse, 1991; McDonough & McDonough, 1990; Nunan, 1989).
7. The relationship between theory and practice is becoming closer (Clarke, 1994; Ellis, 1992; Gaies, 1987; McDonough & McDonough, 1990; Schinke-Llano, 1991; Wagner, 1991; Ur, 1992).
8. There is a feeling of being adrift between two worlds because we don't have a theory to make sense of or to justify our teaching (O'Neill, 1989).
9. There is a trend towards a focus on the learner as an individual. For ex-

ample, learner training and learner development (Brown, 1991; Morley, 1987; Savignon, 1991).

10. There is a trend toward seeing teaching as unique and as having its own skills and knowledge (Freeman & Richards, 1993; Gaies, 1991; Genesee, 1994).
11. There is an increasing awareness that we are a separate discipline and not a branch of linguistics (Bahns, 1990; Grosse, 1991; Nunan, 1989; Richards & Rogers, 1987; Schinke-Llano, 1991; Ur, 1992).
12. We are becoming increasingly professional and more concerned with issues such as the development of credentials, regulations, and entry into the profession (Aitchison, 1993; Ashworth, 1991; Bowers, 1994; Clayton, 1989; Gaies, 1987; Grosse, 1991; Helgesen, 1991; Maley, 1992; O'Neill, 1990; Redfield, 1990; Swales, 1993).
13. There is a trend towards nonnative speakers playing a more important role in the teaching of English. Already most EFL teachers in the world are nonnative speakers (Brown, 1991).
14. There is a trend towards including human issues in our curriculum such as peace and environmental education (Brown, 1991).
15. Psycholinguistics is developing into a more important area for language teachers (Aitchison, 1993).

This list of trends, while well documented in the literature, is not a complete list and perhaps could never be. Rather, it represents one possible set of answers to the questions, where are we now and where are we heading?

The Response to the Survey

My second question is, to what extent do JALT leaders agree or disagree with these trends. Eighty questionnaires were mailed to persons listed as National Officers, N-SIG Officers, and Chapter Officers. As Table 1 shows, surveys were sent to all the National Officers, one representative from each SIG and one representative from each Chapter. The survey was mailed to the person in the SIG or the Chapter who was listed as president or coordinator unless that person was also listed as a National Officers in which case the survey was sent to the program chair. Fifty-five survey questionnaires were returned for a response rate of sixty-nine percent (see Table 1).

Of the 55 valid respondents, the L1 was given as English by 47, Japanese by 6, French by 1, and German by 1. There were 39 male respondents and 16 females. The average years teaching experience was 14.5 years. Some of the respondents teach at more than one type of institution. Five teach at junior high schools, ten at high schools, thirty-four at colleges or universities, eight at commercial language schools, five at home, four at *senmon gakko* (vocational schools), and seven teach in business in-house programs. The most typical respondent

Table 1
Survey response

	mailed	returned invalid	returned	per cent valid
National Officers	8		8	1.0
Appointed Officers	10	1	8	.80
Committee Chairs	7		6	.86
Editors	5		3	.60
JALT 94 Conference	4	1	3	.75
N-SIG Coordinators	10		6	.60
Chapter Officers	36	1	21	.58
Totals	80	3	55	.69

Note: One survey was returned due to an incorrect address and two surveys were returned unanswered.

was an English speaking male working at a college or a university who has about fourteen years teaching experience.

To summarize the trends into one sentence, we are a global (trends 1, 4, and 14) yet local (trends 3 and 5), expanding (trends 2 and 13) group of teachers who are in the process of becoming a profession (trends 8, 10, 11, 12, and 15) and setting our own agenda (trends 6, 7, and 9).

For a breakdown of the answers see Appendix. For a more visual representation of the results, the following pages present the trends seen as bar charts. Percentages have been rounded-off to the nearest whole number and, therefore, the totals are not always exactly 100%.

Additional Comments From Current JALT Leadership

My third question is, what additional trends do JALT leaders perceive? The trends survey contained one open question which asked respondents if there were any trends they felt had been overlooked and if there were, to please write them in a provided space. Following are the additional trends suggested by the individual respondents.

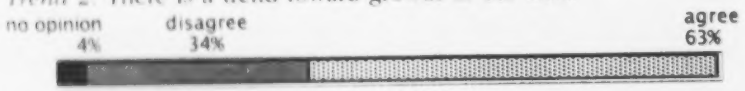
The JALT leadership perceives a trend towards:

1. Students learning more than one foreign language.
2. Students learning languages in the host language country; participating in more overseas English study courses.
3. Studying the similarities in Asian (Japan, Korea, Indonesia, China) learners of English.
4. Language teaching as a business.
5. Using English to teach culture.

BEST COPY AVAILABLE

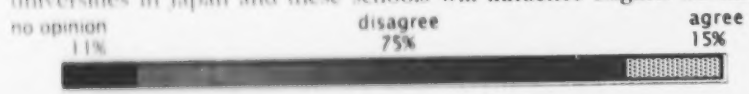
Trend 1. As a knowledge-based field, we are becoming part of the global information network.

Trend 2. There is a trend toward growth in our field.



Trend 3. The new *Monbusho* course guidelines will have a significant impact.

Trend 4. There is a trend towards the establishment of U.S. colleges and universities in Japan and these schools will influence English teaching.

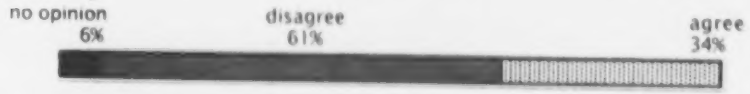


Trend 5. There is a trend toward the JET, AET program having a greater influence on us.

Trend 6. There is a trend towards teachers becoming more interested in classroom-based research.

Trend 7. There is a trend toward the relationship between theory and practice becoming closer.

Trend 8. There is a trend toward feeling lost between two worlds because we don't have a theory to make sense of or to justify our teaching.



Trend 9. There is a trend towards a focus on the learner as an individual, for example learner training and learner development.

Trend 10. There is a trend toward seeing teaching as unique and as having its own skills and knowledge.

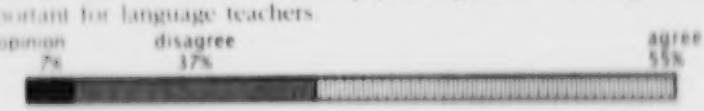
Trend 11. There is an increasing awareness that we are a separate discipline and not a branch of some other discipline, for example linguistics or literature.

Trend 12. There is a trend toward becoming more concerned with the development of teaching credentials.

Trend 13. There is a trend towards English nonnative speakers playing a more important role in the teaching of English. Already most EFL teachers in the world are nonnative speakers.

Trend 14 There is a trend towards including human issues in our curriculum such as peace and environmental education.

Trend 15 There is a trend toward psycholinguistics becoming more important for language teachers.



6. Introducing English into the school curriculum at an earlier age.
7. Children beginning to learn EFL at an earlier age.
8. Junior high school and high school teachers being better speakers of English and therefore graduating better students.
9. Presenting qualitative research under the guise of pseudo-scientific quantitative research.
10. Curriculum evaluation by both teachers and students increasing at the university level.
11. Decentralized authoritarian control.
12. Using part-time, non-tenured, non-professional instructors.
13. Not opting for any one particular method, a trend back toward eclecticism.
14. Looking too hard for trends (especially where *Monbusho* is involved).
15. Teaching English as an international language (EIL).
16. Moving back to grammar.
17. Learning objectives which are more narrow than "communicative competence," e.g., ESP, translation courses, survival courses, English for academic purposes.
18. ESP/EST.
19. Teacher development.
20. Testing.
21. Downplaying the learning of English because the rest of the world is going to learn Japanese.
22. Employing native English speakers more for their market value than for their effectiveness as teachers.
23. Limiting instruction, skills, and content to the elements students will need to become efficient corporate components and heroic consumers.

This list contains some trends that are cryptic and in need of explanation and clarification. In addition, all of these "additional trends" need additional support

such as references from the literature or empirical support from other sources.

In addition to additional trends, some respondents included comments.

Respondent Comments

- There is tremendous variability in the definition of what a language teacher is.
- Teachers should pay more attention to the relationship between CLT and University entrance examination; more study should be done in this field, especially experimental research.
- Some of these trends I wish were happening but are not.
- It is disappointing that many JETs in high schools can't expect universities to change their exam structure. The government should take some measure to introduce aural tests in the national uniform exams.
- I hire part time teachers for universities and I agree that better qualified ESL teachers in terms of degrees are available. An MA or Ph.D. in TESOL plus publications is almost required.
- As a *gaijin* (foreigner) I circled 6 for many of the trends, but I believe that for the Japanese teachers it would be a 1 or a 2.

Discussion

One issue for discussion is who do the respondents of this survey represent and who do they not represent. Not every leader currently active in JALT was sent a survey, and some that were sent a survey chose not to respond. The Japanese leadership is under represented. University teachers are adequately represented, but teachers in other institutions are not, and as a result their insights are lacking. Nevertheless, given these limitations, it is my experience in JALT that the group who responded to this survey do, in fact, reflect the JALT Executive Board as it is presently constituted and in that sense can be said to represent the thinking of the current JALT leadership. It remains for future surveys to document change. For research into trends, both global and in Japan, to continue, it would probably be necessary for teacher groups such as JALT to officially support a larger-scale survey.

Comments on the Survey

This questionnaire survey was not field tested, reports no reliability coefficients, and offers no validation. For these reasons this article should be thought of as an interpretative essay reporting insights rather than research reporting generalizable results. There were two areas of comments by the respondents of the survey. Both areas are concerned with clarity. One is the issue of instrument clarity and the other is the issue of theoretical clarity.

Respondents were asked to circle the number for each question that best showed their opinion. I should have stipulated in the directions to circle only one number because some respondents circled more than one number, thus invalidating their answer. For example, one respondent in answering trend ques-

tion number two, wrote "short term" and circled number 5 and also wrote "long term" and circled number 6. Another indication that my directions were lacking in clarity in that some respondents answered by saying in their opinion the literature indicated one position, but that their experience indicated another position. Some respondents seemed unclear about whether I meant them to answer as if they were in Japan (which they were) or from some other place.

In addition to a lack of clarity on how the survey instructions were phrased, a few respondents sensed a lack of conceptual clarity. One respondent wrote, "What I think or what teachers in general believe?" Another respondent complained that the survey was unclear because it did not specify the reference point as what each teacher personally believes, what native speaker EFL college teachers believe, what non native speaker EFL teachers in Japanese public schools believe, or what the fresh *gaijin* in the JET program believes.

The definition of what a trend is and is not needs to be made clearer. What constitutes a trend? How does one know when a trend exists? Does agreeing with a trend signify that one is acknowledging its existence or expressing approval? These and other questions need to be discussed and taken account of in future trend surveys.

When we look at the trends from the literature listed in this paper and consider the degree of agreement or disagreement of the JALT leadership with these trends, we see that the JALT leadership agrees with most of the trends with some important differences. JALT leadership group does not believe that [...]e colleges and universities are expanding and will have a significant impact on can teaching in Japan (trend 4); it does not yet see the influence of the AET-JET program (trend 5); nor does it feel lost due to the lack of a comprehensive theory that can guide pedagogy (trend 8). In addition, the JALT leadership is not sure about there being a closing of the gap between theory and experience (trend 7). That being the case, it is not surprising that they do not experience psycholinguistics (trend 15) as being particularly helpful.

Given that the topic of this collection of papers is on classroom based research, it would be helpful to look more closely at trend 6, that more teachers are becoming interested in classroom based research. As you will recall, there was 74% agreement with this trend. We will look at the background of this trend as well as some of the issues which are involved.

Classroom Research

The classroom teacher has traditionally been a marginal research participant. For at least the last 20 years, the defining centerpiece of ESL teacher education has been the teaching method taught in the methods course (Gaies, 1981; Numan, 1988). The methods course lacked training in classroom based research and tended to promote a method or methods for teaching (Grosse, 1991). The method (e.g. TPR, Silent Way, CLL) was often promoted by a teaching guru who viewed

BEST COPY AVAILABLE

the teacher as irrelevant at best or a contamination of the process at worst (Gaies, 1981). It was the teaching method that was seen as responsible for learning. The teacher's job was to correctly administer the method and/or stay out of the way of the method as it connected with the students.

The "classic" researcher has special training and special statistical tools with which to conduct research which the classroom teacher almost always does not have. Therefore, the role of the classroom teacher is to receive the research. This sets up a dichotomy of researcher as producer of knowledge and classroom teacher as consumer of knowledge (McDonough & McDonough, 1980). Because research is thus seen as a top-down arrangement coming from the researcher who is specially trained for the job, there is a mismatch between the role of the classroom instructor and the ESL researcher which makes it difficult or impossible for the two parties to communicate.

A new image is currently in the process of being formed for the role of the ESL teacher. This new image is multifaceted in that it includes both classroom practices as well as a research role. Genesee (1984) observed that the title "teacher-researcher" is part of this new image, but a mixed metaphor. On the "teacher" side of the metaphor, Genesee agreed that more teachers are currently involved in research than before and furthermore that this research is increasingly classroom based. On the "researcher" side, he expressed concern that teachers may be judged by criteria appropriate to researchers but not teachers. Allwright (1984) responded by suggesting that Genesee would be correct only if he accepts the classical definition of the researcher. Allwright pointed to the classroom and teacher pedagogical practices as the focus of an image of research more suited to teachers and the language classroom. McDonough & McDonough (1980) called this type of classroom research bottom-up research as opposed to the more traditional top-down research model.

Gaies (1981) articulated an image of the teacher as manager because in his view teaching is decision making. It is the nature and process of teacher decisions that ought to be looked at carefully in classroom research, especially as they promote change in classroom pedagogy and practices. Nunan (1989, 1990) said teachers should be managers and researchers of their own teaching practices. To accomplish this, teachers need special training in classroom research. Appropriate forms of research include action research and ethnographic research both of which look closely at what actually happens in classrooms as opposed to what teachers believe happens. This view is shared by Greene (1981, p. 44) who stated that a major challenge to teacher education is the "gap between what teachers and researchers think happens in the classroom and what actually goes on." A major concern underlying this discussion of teacher as researcher is how we as classroom teachers initiate, promote, and manage change and innovation in our classrooms, in our institutions, and ultimately in the larger society.

Conclusion

This paper began with the image of a train. The train was moving in unfamiliar territory and the question was asked, where are we now and where are we going? I will continue to use the pronoun "we" to indicate that the answer to this question affects the entire ESL/EFL teaching profession.

The answers to this survey seem contradictory. On the one hand, we believe teacher training and credentials are important and we are interested in classroom research, but on the other hand we don't feel a narrowing of the gap between practice and research. We believe non-native speaker teachers are becoming more important, but we suspect many non-native speaker teachers don't share this opinion. We think the new *Monbusho* course guidelines will significantly impact us, but we don't think the JET/AET program is having much influence.

The tentative conclusion reached earlier in this paper was that we are a global-local group of teachers in the process of becoming a profession and setting our agenda. It is this last issue, setting our agenda, where I sense we are not clear and which I believe gives rise to our contradictory responses. At the moment we are getting ready for the agenda task by deciding what issues we think are important and preparing ourselves for the research task ahead. We are not sure though where we want to go or how to get there, but we sense that as classroom teachers it is up to us to make those decisions and preparations and not ask others to do it for us. Small wonder that we are a bit confused.

You are now back on the train. You finish your class and look out the window as the train approaches the station. As the train pulls in you notice a sign which announces the name of the station. The sign says, Welcome To Getting Ready. You get off the train and think to yourself, "This must be the place."

References

Allwright, D. (1994). Teachers as researchers: Un-making the metaphor? *TESOL Matters 4* (5).

Aiga, Y. (1990). Is Japanese English education changing? *Cross Currents 17*, 139–143.

Aitchison, J. (1993). Birds, bees, and switches: Psycholinguistic issues 1963–2013? *ELTJ 47*, 107–116.

Alatis, J. (1987). The growth of professionalism in TESOL: Challenges and prospects for the future. *TESOL Quarterly, 21*, 9–19.

Ashworth, M. (1991). Internationalism and our "strenuous family." *TESOL Quarterly, 25*, 231–245.

Bahns, J. (1993). Consultant not initiator: The role of the applied SLA researcher. *ELTJ 44*, 110–119.

Bowers, R. (1994). English 2000. *Practical English Teacher, 14* (3), 31–35.

Brown, H. D. (1991). TESOL at twenty-five: What are the issues? *TESOL Quarterly, 25*, 245–260.

Carter, N., Goold, R., & Madeley, C. (1993). The new Monbusho guidelines, part two. *The Language Teacher, 17* (11), 3–7.

Clarke, M. (1994). The dysfunctions of the theory/practice discourse. *TESOL Quarterly, 28*, 9–26.

Clayton, T. (1989). International teaching of English to speakers of other languages: Where is our profession going? *Cross Currents, 16*, 55–61.

Cominos, A. (1992). Managing change in foreign language education: Interview with Minoru Wada. *The Language Teacher, 17* (11), 3–7.

Ellis, R. (1992). *Second language acquisition and language pedagogy*. Clevedon: Multilingual Matters.

Fanselow, J. (1987). TESOL and the future. *TESOL Newsletter, XXI* (1), 9.

Freeman, D. & Richards, J. (1993). Conceptions of teaching and the education of second language teachers. *TESOL Quarterly, 27*, 193–216.

Gaies, S. (1987). Research in TESOL: Romance, precision, and reality. *TESOL Newsletter, 21* (2), 21–23.

Gaies, S. (1991). English language teaching in the 1980s: How will the teacher fit in? *JALT Journal, 13*, 7–21.

Genesee, F. (1994). Teachers-as-researchers: Mixed metaphor or mixed blessing? *TESOL Matters, 4* (2), 5.

Goold, R., Madeley, C., & Carter, N. (1993). The new Monbusho guidelines. *The Language Teacher, 17* (6), 3–5.

Grosse, C. (1991). The TESOL methods course. *TESOL Quarterly, 25*, 29–40.

Hamel, G. & Prahalad, C. K. (1994, September 5). Seeing the future first. *Fortune*, pp. 65–68.

Helgesen, M. (1991). Teaching in Japan. *TESOL Journal, 1* (1), 8.

Leedham, M. (1991). The JET Programme: Sticking plaster for English Education? *The Language Teacher, 15* (6), 21–25.

Maley, A. (1992). An open letter to the profession. *ELTJ, 46*, 96–99.

McDonough, J. & McDonough, S. (1980). What's the use of research? *ELTJ, 44*, 102–109.

Morley, J. (1987). Current directions in second language learning and teaching: A state-of-the-art synopsis. *TESOL Newsletter, 21* (2), 16–20.

Nunan, D. (1989). Second language teacher education: Present trends and future prospects. In C. N. Candlin & T. F. McNamara (Eds.), *Language and community* (pp. 143–154). Macquarie University, Sydney: National Centre for English Language Teaching and Research.

Nunan, D. (1990). Hidden voices: Insiders' perspectives on classroom interaction. In K. Bailey & D. Nunan (Eds.), *Voices from the language classroom: Qualitative research in second language education* (pp. 41–56). Cambridge: Cambridge University Press.

O'Neill, R. (1980). In praise of intellectual obscenity or: the confessions of an embarrassed eclectic. *Cross Currents, 15* (2), 1–10.

O'Neill, R. (1980). What did Isadora Duncan and Einstein have in common? *Cross Currents, 17*, 84–87.

Redfield, M. (1988). Japanese university hiring practices and the masters in TEFL. *Cross Currents, 17*, 89–91.

Richards, J. & Rogers, T. (1987). Through the looking glass: Trends and directions in language teaching. *RELC Journal, 18* (2), 45–53.

Savignon, S. (1991). Communicative language teaching: State of the art. *TESOL Quar-*

terly, 25, 261–277.
Schinke-Llano, L. (1991). TESOL: Ironies of the field. *TESOL Journal, 1*, (2), 5–6.
Shang-Ikeda, S. (1992). Interview: Robert Jupp. *The Language Teacher, 16* (10), 31–35.
Swales, J. (1988). 20 years of the TESOL Quarterly. *TESOL Quarterly, 22*, 151–163.
Swales, J. (1993). The English language and its teachers. *ELTJ, 47*, 283–291.
Ur, P. (1992). Teacher learning. *ELTJ, 46*, 56–61.
Wada, M. (1994). Team teaching and the revised course of study. In M. Wada & A. Cominos (Eds.), *Studies in team teaching*. Tokyo: Kenkyusha.
Wada, M. (1995, June 26). Team teaching becomes fact of classroom life. *The Daily Yomiuri*, p. 9.
Wada, M., & Cominos, A. (1994). *Studies in team teaching*. Tokyo: Kenkyusha.
Wadden, P. (1994). College curricula and the foreign language teacher: A forecast for the late nineties. *The Language Teacher, 18*, (7), 32–35.
Wagner, J. (1991). Innovation in foreign language teaching. In R. Phillipson, E. Kellerman, L. Selinker, M. Sharwood Smith, and M. Swain (Eds.), *Foreign/Second language pedagogy research* (pp. 288–306). Clevedon: Multilingual Matters.
Widdowson, H. G. (1986). Forty years on. *ELTJ, 40*, 265–269.
Widdowson, H. G. (1992). ELT and EL teachers. *ELTJ, 46*, 333–339.
Yukawa, E. (1992). Team teaching and changes in teaching routines. *The Language Teacher, 16* (11), 9.

Appendix: Survey Results

	Percent No opinion	Percent Tend to disagree			Positive Response Tend to agree		
Trend	0	1	2	3	4	5	6
Trend 1	11	02	07	15	15	30	20
Trend 2	04	02	1	22	22	29	12
Trend 3	13	05	07	15	29	22	09
Trend 4	11	22	38	15	11	04	00
Trend 5	02	13	25	20	15	22	04
Trend 6	04	04	07	11	20	39	15
Trend 7	04	00	17	26	22	26	06
Trend 8	06	22	22	17	13	19	02
Trend 9	04	00	06	06	28	39	17
Trend 10	08	00	13	12	17	33	17
Trend 11	08	04	11	04	25	26	23
Trend 12	07	00	06	02	15	43	28
Trend 13	06	00	04	10	24	39	18
Trend 14	06	00	09	06	26	35	19
Trend 15	07	04	09	24	31	17	07

SECTION II

Getting Started

Chapter 3

First Things First: Writing The Research Proposal

Martin White

International Communications Program, AIG, KK

Diane White

Tokyo Women's Christian University

A common way to do research is to plunge straight in, collect a vast amount of data, and then: disaster! You quickly panic as you struggle to find a way to process all the collected information.

It is far better to approach research in a systematic way. No builder would start construction without having first drawn up a set of architectural plans. In the same way, a research proposal is a modest document of one to four pages which shows concisely what the research topic is and how you are going to carry out the research. Having a properly written research proposal helps ensure that all parts of the research project have been carefully thought about before the research begins. Without careful consideration beforehand, you run the risk of coming up against serious problems that can result in a lot of time wasting and frustration.

Sometimes a research proposal is compulsory. That is the case if you are doing research in the hope of getting an academic qualification such as a Masters or Doctoral degree. Research is often funded by foundations or government agencies. Such bodies will require a research proposal before approving the research or granting any funds. But even when it is not mandatory, it still makes sense to do a research proposal as part of the process of planning a

research project. Not only will this alert you to potential problems, but you will then have a handy document which you can use when you consult other, more experienced people for advice on how to solve those problems. With your proposal to work from, your more experienced colleagues are able to give you better advice on frequently encountered difficulties such as analyzing statistics by computer, and choosing suitable subjects.

For several reasons then, writing a research proposal can be an important and useful step in the process of planning and implementing a research study. But, what exactly is a research proposal and how does one go about writing one? What follows is an outline to help in the writing of a proposal. Each step is explained and an example taken from the proposal written by one of the authors (see Appendix). The completed research study was briefly discussed in Ellis (1994).

The form that this research proposal outline takes is similar to that of the final document written after the data has been collected and analyzed. This will help save time and make the job of writing the final report easier.

Title

A. Introduction
   1. Context—description of topic or problem
   2. Review of literature
   3. Research questions or hypotheses, or both

B. Methodology
   1. Subjects
   2. Design
   3. Data collection

C. Analysis
   1. Data organization
   2. Statistical procedures

D. Significance of Study

Title

The first thing you need is a working title for your research. Research titles have run the gamut from lengthy, arcane stuffy titles to instant attention-getting quirky titles such as "Say it Again Sam" or "The Chicago Which Hunt." Our advice is to steer a middle way. Have a title which gives the reader a clear idea of the topic but which is also brief. One way to do this is to use a short statement followed by a colon. After the colon can be a lengthier delineation of the topic. Thus is the approach taken by White (1992) in deciding on the title, "Teachers' Questions: Form, Function and Interaction—A Study of Two Teachers."

Introduction

The research proposal begins with an introduction. Usually this is not labeled specifically "Introduction." This convention is also followed by journal articles. Check accounts of research in such journals as *TESOL Quarterly* and *JALT Journal*. As indicated in the above outline, the introduction has three parts. The first part gives the reader an idea of the focus of the research. The second part attempts to link the research issue with existing information. Having reviewed the existing literature—the current body of information relevant to your topic—you need to establish why your research is important and timely. One way is to show that there is a lack of information relating to your topic or alternatively, that you can add a significant and different perspective. Having established a need for your research, you can then state your research questions or research hypotheses, or both.

Nunan (1992) stated that the minimum requirements for an activity to be considered research are: (a) a question; (b) data; and (c) analysis and interpretation. You cannot begin to plan or carry out research until you have a research question. The first step is to find a topic or area that interests you. This can come from your experience as a classroom teacher or language learner or from your own reading. Reading widely in academic journals can give you ideas for research. If you are interested in second language acquisition it is a good idea to read accounts of how children learn their mother tongue, or journals in fields other than language learning. For one piece of research we read journals dealing with social psychology and found ideas for a study dealing with the different factors that cause people to adjust their speech for different partners. If you are really in pursuit of ideas for research then the conclusions of doctoral theses are good hunting grounds because it is customary in the closing of a thesis to mention questions that the current study did not answer and to give recommendations for further study.

Having found an area that interests you, you will need to form a research question. Good research questions meet two criteria: (a) the question merits answering and is worth seeking an answer for, and (b) the question can, in fact, be answered.

In the research on teachers' questions, we started our proposal with the following statement to define what the research was about: "The kinds of questions teachers ask in classrooms has increasingly attracted the attention of researchers in recent years." We then did a brief survey of the existing research dealing with teachers' questions.

The Literature Review

The literature review has several functions: (a) it gives the background to your proposed research, (b) it shows that you are competent in that you have an understanding of what is currently known, and (c) it provides a reason for your research.

The review section usually ends with the author discussing the implications arising from the existing literature. Here is the place to point out gaps in existing knowledge on the subject, or the need to look at the issues from a fresh perspective. In this way, you provide justification for your proposed study, and demonstrate its importance.

In our proposal, the review of the literature pointed out that recent research on teachers' questions had moved towards experimental and quantitative research with reports containing a lot of statistical analysis but fewer examples of actual classroom discourse when compared to previous research in L1 settings. We mentioned that the study proposed "to fill this gap by not only analyzing the frequency of questions according to form and function but also. . . includ[ing] the additional dimension of discourse analysis."

Questions or Hypotheses

This section states your carefully thought up and worded hypothesis. In our research on teachers' questions we asked three questions:

1. What type of questions do teachers ask?
2. Which types of questions result in the longest student response?
3. Will there be any difference in the types of questions and responses according to the proficiency level of the class?

In addition to research questions, we also stated the object of our research in the form of a hypothesis. If your purpose is to acquire information—for example, the order in which second language learners learn morphemes—then the research area you will tackle is put as a research question. If, however, you are testing a theory, then you should state the object of your research as a hypothesis. The difference is that the answer to a hypothesis is a simple yes or no. The theory was either proved or it was not. With a research question, however, you can have a much wider range of answers. In our paper we included three hypotheses:

1. Teachers will ask more display questions than referential questions.
2. Learner responses to referential questions will be longer than the response to display questions.
3. A greater number of referential questions will be accompanied by a greater number of confirmation checks and clarification requests by the teacher.

Method

In this section of the proposal you include a description of how the research will be carried out and the method you will adopt to answer your research questions or hypotheses. You include who the subjects will be, the kind of research it is and how you will collect the data.

One test of a good methodology section is if after reading it somebody could proceed and carry out the research based solely on the information you have given. A straightforward way to tackle this section is to have three subcategories: (a) subjects, (b) design, and (c) data collection.

In our proposal on questions, we started this section by explaining that "the subjects for this study will be 43 adult students attending the in-company language training program of a major multi-national company." Next we profiled the teachers who would be the focus of the study. At this point we also explained the basis for choosing these two teachers. "The two teachers were selected on the basis of similar sex, nationality, and teaching experience. They also had access to classes at comparative levels and could teach the same lesson material to their classes."

Design

This section is the place to describe the type of research you are going to carry out. In our report we handled it this way: "This is intended to be descriptive research and will use naturally existing classes for the purpose of data collection. It will be a one-shot design in that each class will be observed only once."

Data Collection

In this section should include a detailed and accurate account of how you will go about collecting the data you need. In addition, this section should mention what equipment or materials will be used. Questionnaires are often listed here if the research in question is of a survey nature. The information in this section should be so clear that the reader can not only visualize the process of how your are going to do your research but would also have sufficient information to replicate your research. It should also be clear to the reader that the steps you have outlined are sufficient to answer the research questions mentioned. In our research proposal we mentioned that we would take four audio recordings of one hour each, and we listed the teachers and the classes that would be recorded. We also mentioned the newspaper article and the material that was to be used in the classes.

Analysis

This part of the proposal has typically two parts: (a) data organization, and (b) statistical procedures. This section focuses on the question of what you are going to do with all the data when you have collected it. It can be frightening to collect thousands of pieces of data and not know what to do with the information. Thinking in advance and putting something down in writing before you start can save a lot of frustration and wasted effort. For our paper on teachers' questions we tracked down, in the literature, various schemes for categorizing questions. The resulting classification scheme, a taxonomy, was largely what we had read in books and journals but was also modified to better suit the aims of the research. If you are going to use charts and tables you can also mention them at this point.

Statistical Procedures

This part can be the most daunting of the whole research project. Luckily, if you need it, you can get help to steer you through the maze of deciding whether you need to use ANOVA, Pearson Correlation, or Chi-square. We found what statistics would help analyze our data to prove or disprove our hypotheses and were able to write, "Statistical frequencies obtained will involve 2-way Chi-square." Not all data will need a computer to process the numbers. For our paper on teachers' questions we did a lot of counting by hand. That involved adding up the number of times a student or teacher took a turn in speaking, and calculating the mean number of words in each student response.

Significance

This section is the place for you to state the importance of your research study. In it you explain the contribution it will make to linguistic science or whatever field you have chosen to write about. In the proposal on teachers' questions, we set about the task in this way.

> To the best of our knowledge no study has been made using Well's Discourse Role and Commodity Network for an analysis of the exchange structure of classroom discourse in an adult EFL setting. The effect of different proficiency levels has also not yet been researched. The proposed study outlined above is intended to contribute to a greater understanding of the role of teachers' questions and to provide more illustrative samples of actual classroom discourse.

It is said that there is nothing so practical as a good theory. The research you are proposing to carry out may have the potential to change peoples' lives—or at least how they teach and learn languages. If the results of your research have this potential here is the place to make the announcement.

Reviewing the Proposal—Potential Problems

When you have written your research proposal, you should review it, possibly with the help of a colleague. One of the authors, who teaches research writing classes at a Japanese university, has found that the biggest problems that novice researchers have is defining the research question. As Nunan (1995, personal communication) has pointed out, this can also be a problem for experienced researchers. The problem is that often would-be researchers are unclear about knowing what they actually want to study. It is crucial that a beginning researcher has a clear understanding of what they want to research. If you have carefully thought out what you want to study and defined this in a suitable research question then the research has the greatest chance of going very smoothly. Read over your research proposal and make sure that it is clear what your study actually proposes to do. Make sure too, that the topic is worth investigating. Next check that the topic is sufficiently focused. Another major

problem with many initial proposals is that the scope of the research is too broad. This leads to problems of data collection and analysis and if it is too broad then it may be impossible for the research to lead to any conclusive findings. Our advice would be to keep the focus of enquiry relatively simple.

When reviewing the proposal also examine carefully the choice of subjects and the procedures that will be used. Make sure that you have carefully considered the relevant variables. For example, if you are comparing two groups and one group performs better using your material, can you safely say it was solely due to the effectiveness of the materials? Perhaps the group that did better was a higher level group or those students usually perform better since their teacher is more enthusiastic and they are therefore more motivated to learn English. Also make sure you can justify your use of subjects. It is not enough to reply that it is just convenient to use this set of students or that it is easier to get permission from a certain teacher.

Another problem frequently encountered is the difficulty of finding relevant background material. The following journals are a good source of articles that cover teaching, learning, and applied linguistics. *Applied Linguistics, ELT Journal, English Teaching Forum, JALT Journal, Language Learning, Studies in Second Language Learning, System, TESOL Quarterly.*

Researchers interested in using the computer to do research can access ERIC (Educational Resources Information Center). ERIC is an abstract service that lists important journal articles and conference papers according to various categories.

Conclusion

Writing a research proposal is a preparatory stage in the cycle of doing a research study. Seliger and Shohamy (1989) stated that there are four phases in the preparatory stage of any research study.

Phase 1 Formulating the general question. Possible sources are the experience and interests of the teacher/researcher, other research on second language acquisition or sources outside second language acquisition.

Phase 2 Focusing the question. Is the question both feasible and important?

Phase 3 Deciding on an objective. Description of research procedures and formulating research questions.

Phase 4 Formulating the research plan.

The research proposal is really a plan, a step-by-step guide to doing research. It will help ensure that you have given sufficient thought to all the crucial areas of research, such as, the basic research design, data collection and analysis of the results. It will also help you to see and avoid many potential problems. In addition to helping give shape to the planning process, it also makes the final

writing up of the research easier as you already have the main headings for each part of the final research report. Once you have that done, you are ready for publication. In Japan two widely read journals that publish research are the *JALT Journal* and *The Language Teacher*. More and more teachers are carrying out their own research, particularly classroom-centered research. By following a step-by-step plan such as that outlined above even those not trained in research or statistical procedures could start a research project and make a valuable contribution.

References

Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.

Nunan, D. (1992). *Research methods in language learning*. Cambridge: Cambridge University Press.

Seliger, H., & Shohamy, E. (1989). *Second language research methods*. Oxford: Oxford University Press.

White, M. (1992). Teachers' questions: Form, function, and interaction—A study of two teachers. Unpublished paper, Tokyo: Temple University Japan.

Appendix: A Research Proposal

Title

Teachers' Questions: Form, Function and Interaction—A Study of Two Teachers

Introduction

The kinds of questions teachers ask in classrooms has increasingly attracted the attention of researchers in recent years. Studies of teachers' questions (Long and Sato, 1983; Brock, 1986; Long & Crookes, 1986) have analyzed the frequency of question types according to function, and in particular have sought to discover whether teachers ask more display questions than referential questions. The results indicate that in ESL classrooms teachers ask too many questions (White & Lightbown, 1984) and that these questions are typically display questions (Long & Sato, 1983). Studies in teachers' questions have also concentrated on identifying the effects of the use of comprehension checks, confirmation checks, and clarification requests. Such modifications to the interactional structure of conversation, it is claimed, facilitate second language acquisition. Earlier studies of teachers' questions featured school classrooms in an L1 setting. Studies such as Barnes, (1969) and Mehan, (1978) are informative and illustrated with extracts of classroom discourse. Later research, however, focusing on the ESL classroom has moved in the direction of quantitative research and contains almost no examples of recorded classroom discourse. In addition, studies such as Long and Sato (1983) and Brock (1986) have been experimental in design and have grouped students especially for the experiment or given advice to the teachers on how to teach the lesson.

The proposed study intends to fill this gap by not only analyzing the frequency of questions according to form and function, but also by including the additional dimension of discourse analysis.

The following research questions will be addressed:

1. What types of questions do teachers ask?
2. Which types of questions result in the longest student response?
3. Will there be any difference in the types of questions and responses according to the proficiency level of the class?

In addition, the following three research hypotheses, based on Brock (1986) will be tested:

1. Teachers will ask more display questions than referential questions.
2. Learner responses to referential questions will be longer and syntactically more complex than the response to display questions.
3. A greater number of referential questions will be accompanied by a greater number of confirmation checks and clarification requests by the teacher.

Method

Subjects

The subjects for this study will be 43 adult students attending an in-company language training program at a major Japanese multinational company, for four hours a week. In terms of proficiency as measured by TOEIC scores the higher class has an average score of 711 and the lower level class has an average TOEIC score of 424. All the students have had at least five years of learning English either at school, university, or in company classes. Also serving as subjects for the study are two teachers, both female and British. Both teachers, have the RSA Certificate in TEFL and one teacher also has the RSA Diploma and a Postgraduate Diploma in EFL and is presently writing a thesis to complete her master's. The other teacher is in the process of earning the RSA Diploma in EFL. The teachers were selected on the basis of similar sex, nationality, and teaching experience. They also had access to classes at comparative levels and could teach the same lesson material to their classes. The students will be grouped in their already existing classes and taught by their normally scheduled teacher.

Design

This is intended to be descriptive research and will use naturally existing classes for the purpose of data collection. It will be a one-shot design in that each class will be observed only once.

Data Collection

Audio recordings of one hour duration will be made from four classes in total: two from a low proficiency class and two from a high proficiency class. Both teachers will teach one low proficiency class and high proficiency class. The teachers will use the same material for classes of a similar level. The higher level class will use an authentic, unsimplified article on illegitimacy taken from the British newspaper, *The Daily Telegraph*. The material used for the lower proficiency classes will be Unit 11 of *Coast to*

Coast (Harmer, 1987), a general English textbook. The teachers will be provided with a small professional tape recorder to place on the table. These recordings will be transcribed and coded for the frequency of question types.

Analysis

Taxonomy for the form of teacher questions:

Wh-questions, Yes/No questions, Alternative questions, Rhetorical Function taxonomy (modified version of Long and Sato, 1983).
Display questions—form-focused/content focused/repetitions.
Referential questions—open/closed/repetitions.
Rhetorical/expressive/repetitions.
Comprehension checks, confirmation checks, clarification requests.

The number of turns per student and teacher and mean number of words for student responses after given type of question will be calculated. Statistical analysis of frequencies obtained will involve 2-way Chi-square. To examine the discourse role of questions, the system shown in Wells (1979) will be used.

Significance of study

To the best of our knowledge no study has been made using Wells's discourse role and Commodity network for an analysis of the exchange structure of classroom discourse in an adult EFL setting. The effect of different proficiency levels on teacher's questioning behavior has also not yet been researched. The proposed study outlined above is intended to contribute to a greater understanding of the role of teachers' questions and to provide more illustrative samples of actual classroom discourse. According to Long and Sato (1983), their results show that communicative use of the target language makes up only a minor part of typical classroom activities. Is the ESL classroom really so bereft of possibilities for genuine communication? It is the purpose of the study outlined in this proposal to investigate the truth of such an assertion.

References to the Appendix

Barnes, D., Britton, J., & Torbe, M. (1969). *Language, the learner and the school*. Middlesex, England: Penguin Books.

Brock, C. (1986). The effects of referential questions on ESL classroom discourse. *TESOL Quarterly 20*, 47–58.

Harmer, J., & Surguine, H. (1987). *Coast to coast: Book 1*. Harlow: Longman.

Long, M., & Crookes, G. (1986). Intervention points in second language classroom processes. *University of Hawaii Working Paper Series*, 83–112.

Long, M., & Sato, C. (1983). Classroom foreigner talk discourse: Forms and functions of teachers' questions. In H. Seliger and M. Long (Eds.), *Classroom oriented research in second language acquisition*. (pp. 268–85). Rowley, MA: Newbury House.

Mehan, H. (1978). "What time is it, Denise": Asking known information questions in classroom discourse. *Theory Into Practice 18* (4), 285–294.

Wells, G., Montgomery, M., & Maclure, M. (1979). Adult-child discourse: Outline of a model of analysis. *Journal of Pragmatics 3*, 337–380.

White, L., & Lightbown, J. (1984). Asking and answering in ESL classes. *Canadian Modern Language Review 40*, 228–244.

Chapter 4

Literature Reviews: Obtaining Perspective

Paul Riley

Oxford University Press

"Would you tell me, please, which way I ought to go from here?"
"That depends a good deal on where you want to get to," said the Cat.
"I don't much care where—" said Alice.
"Then it doesn't matter which way you go," said the Cat.
– Lewis Carroll, *Alice in Wonderland*

The question central to all research is, "Which way should I go from here?" Response to this question requires background knowledge to broaden the field of inquiry and then further research to narrow the parameters and frame the research question. This paper looks at the literature review—the researcher's attempt at gathering the necessary background information to figure out where one is and, subsequently, where one is going. It is aimed at teachers with little or no research experience and endeavors to give step-by-step guidelines for conducting and using a review of Second Language Acquisition (SLA) literature.

What is a Literature Review?

A literature review or literature search is essential for any research project. It usually takes place after the researcher has decided on an area of research interest, but before the research question has been formulated. Often there is a second and even third review later in the research process to expand or tighten up the scope of the project. The literature review entails the researcher locating, in journals and

books, research articles relevant to his or her research interests. After collecting a wide and thorough sample of pertinent articles, the researcher must read, synthesize and report on the information uncovered during the survey.

Brown (1988, p. 46) wrote that a literature review should "provide the background or rationale for the study, a demonstration of how previous research is related to the study, and a framework for viewing the study." Nunan (1992, p. 216) agreed that, "the literature review, if carried out systematically, will acquaint you with previous work in the field, and should alert you to potential pitfalls in the chosen area." He added the caution that, "you may come across a study which answers the very question you are proposing to investigate." Seliger and Shohamy (1989, p. 65) referred to literature reviews as "contextualizing the research." They argued that a survey of the literature is undertaken for the seemingly contradictory purposes of broadening the perspective of the project and narrowing down the topic, "because in research there is a need to both expand the perspective and to narrow it down in order to arrive at a workable research question."

Although the terms and definitions differ slightly from researcher to researcher, there is agreement that a review of the literature is necessary for meaningful research. In fact, without a proper survey of the literature, there could be trouble formulating the research question. The intention of a review is not to boldly go where no one has gone before, but to follow the paths others have pioneered before striking out into the uncharted wilderness. Undertaken in the pre-research question stage, it helps to build a foundation for the project and ensures that the researcher does not reinvent the wheel. Performed in the middle of a project, it broadens the researcher's perspective, and suggests new avenues for inquiry. Conversely, late in the research process it helps to focus the evaluation of data, and further clarify the research question.

How to Begin

It is not my intention to grapple with probably the most difficult stage of research: research design and formulating the research question. Suffice it to say, the first thing to be determined, at least in a general sense, is the research idea. This is usually a difficult task for the novice and requires curiosity, observation, and reading on the part of the potential researcher (see Griffee, 1994). But, as Seliger and Shohamy (1989) pointed out, it becomes easier with experience because "research is cyclical. It is an ongoing activity which is never totally completed because each piece of research raises additional questions for more research."

When the research idea has been decided, the first thing to be done is a general literature survey to determine the history (or lack thereof) of the proposed topic. If similar studies have been carried out, the researcher should use the findings to determine the present state of the field and decide if there are any unanswered questions. However, if there have been no similar studies, the researcher must ascertain the reasons why and determine the feasibility of conducting such re-

search. At this point the researcher will need to define concepts, determine the type of study (e.g. heuristic or deductive), and formulate the research question.

Obtaining References

There are many ways to begin a general survey. The best is to speak with an author or researcher who has primary knowledge in the area of interest. These people gather at large language conferences, such as the JALT National Convention or the TESOL Annual Convention, and present papers in their special fields. It is usually possible to attend these presentations and with persistence meet the speaker and ask for advice. As well, these speakers often distribute bibliographies or recommend articles for further study.

Another source of material that can be found at these conventions are the publishers themselves—major publishers will all have displays manned by knowledgeable representatives. Although biased towards their own products, the representatives will suggest appropriate books, which have useful information and unbiased bibliographies. A third option is to ask other teachers in attendance at the conference for help. By asking enough people, the enthusiastic novice is sure to find someone who can point her or him in the right direction.

Technology offers the opportunity to communicate with fellow researchers through the Internet. Anyone with access to a computer with a modem can send and receive information from anywhere in the world quickly and economically. According to Busch (1994, p. 17) a good place to contact others in the SLA field is TESL-L, a computer bulletin board sponsored by the International English Institute of Hunter College, which is located in New York City. Over 1,400 ESL teachers and academics from 48 countries use the bulletin board to communicate on topics concerning second language acquisition. TESL-L can be joined by sending the message "Sub TESL-L <first name last name>" to <listserv@cunyvm.bitnet> or <listserv@cunyvm.cuny.edu>.

A different way to get started is by reading a general applied linguistics book (they usually contain massive literature reviews), that gives an overview of the issues in the second language acquisition field. A couple of the best are *The Study of Second Language Acquisition*, Ellis (1994) and *An Introduction to Second Language Research*, Larsen-Freemen and Long (1991). These are comprehensive studies of the SLA field, and are invaluable as an introduction and research reference. As well, both have extensive bibliographies.

Of course, the most tried and true method of obtaining suitable references is to go to a library and start digging around. Unfortunately, because teachers of English as a foreign language (EFL) are often living and working in places where suitable libraries are scarce, this kind of research can be difficult or impossible. One option is to approach university libraries and ask for permission to use their facilities. Usually, international universities or local campuses of foreign (especially American) universities have the resources for second language research and will allow

use of their library for a fee. Privileges are restricted, however, and often a letter of introduction from a home library or a guarantor is required as part of the application. For more information on how to gain library access, read "Using Japanese Libraries to Do Second Language Research" (Busch, 1994), or call university libraries directly to find out about their visitor policies.

Those lucky enough to have a good library nearby should first check if the library is connected to ERIC (the Educational Resources Information Clearinghouse). ERIC is a computer database which lists documents, journals, and serial publications on different topics related to education. Researchers using ERIC enter in a field of descriptors—defining the parameters of the search—and receive a list, including abstracts, of citations that meet those conditions. Any documents or journal articles that seem relevant can be ordered for a fee. For more details, read "Research in Japan" (Baskin, 1994), which gives an excellent overview of using ERIC and ordering documents from within Japan.

If ERIC is not an option, the best place to start is recent copies of any of the major journals in the second language field: *Applied Linguistics*, *Language Learning*, *Second Language Research*, *Studies in Second Language Acquisition*, and *TESOL Quarterly*. These journals are all widely available, and many print an index—usually in the first or last issue of the year—of all articles received in the previous 12 months. Scanning the titles in the index should offer a number of possibilities which can be assessed by finding the article and reading the abstract.

Extracting Relevant Information

Though there are many ways to obtain references, there is only one way to extract the relevant information, and that is to read. After collecting a stack of articles or list of references, the researcher must separate the wheat from the chaff by reading the abstracts to decide if the material is germane to the research interest and worth reading. Again, the novice may find this difficult, but as one's knowledge of SLA broadens, articles can be assessed more quickly. Criteria to consider when determining the importance of material include the relationship of the material to the proposed research, the source of the material, and recency of publication. Of course, articles with a direct relationship to the proposed research will be more important than those with only a secondary or tertiary connection. Likewise, primary information from well known sources is preferable to second-hand reports or information from unknown or unpublished sources. Finally, researchers should try to review the most up-to-date literature available. It is necessary to obtain the most recent material to ensure coverage of the latest developments in the area of interest.

It is a good idea to devise a note-taking system to organize the reading of the literature. Seliger and Shohamy recommend using a computer database or index cards, and I prefer a small notebook, but the information that needs to be recorded

remains constant regardless of the medium. For each article the researcher must copy down bibliographical information, and should identify and summarize the main arguments or findings. It is important to point out the relevance of the material to the researcher's own interests and write down any ideas on how the information fits into a conceptual framework of the field. The researcher may also want to identify and mark possible quotations to be used when reporting on the review. It is helpful to organize these summaries in some order: alphabetic, by importance, or chronologically. Irrelevant materials should be discarded.

Reading the articles is time consuming and can become overwhelming, as each article offers further areas for inquiry and the research process becomes a series of unending trails. This broadening of perspective often leads into other disciplines (Sociology, Psychology, Education, Linguistics). Brumfit (1995, p. 37) encouraged that tendency: "Research needs to be interdisciplinary, because our reactions to the people with whom we work cannot be constrained by single disciplinary perspectives." The question then becomes, "How much is enough?"

Literature reviews are conducted for different purposes, each with specific requirements. If the researcher is writing for a pedagogical journal like *The Language Teacher*, a brief background check with no more than a half dozen to a dozen sources is usually sufficient. However, research oriented periodicals like the *JALT Journal*, *TESOL Quarterly*, or *Applied Linguistics* require more substantial investigation. Bibliographies in these journals are often a few pages long. The number of relevant sources should act as a guide to the inexperienced researcher. Too many sources may mean the topic is too broad and needs to be narrowed down. Conversely, too few references may indicate a need to expand the scope of the research, or the need to work harder to dig up relevant material.

Writing the Review

As stated earlier, while reading and summarizing the various articles, the researcher needs to simultaneously organize the extracted information, and try to fit it into its place in the survey. A literature review should not be a discussion of one article after another—one paragraph per study. Instead, the researcher should try to develop a framework in which a conceptual order is imposed. This conceptual framework should identify and define the important issues or variables in the field, and the researcher should organize the review to discuss each one in the context of the literature. What should emerge is a clear picture of the history and current state of the research area, what Brown (1988, p. 46) called, "a framework for viewing the study." He went on to add that "an author who is broaching an unexplored area of research cannot cite previous works [but] should at least explain the route by which this new area was reached."

Once the past and present have been discussed, it is time to consider the future. The researcher must state the rationale of his or her own research

project, and consider the significance of the research for language pedagogy. This rationale could contain an evaluation of earlier research, and should include a statement of purpose to justify the proposed research project. Often here is where the survey will end.

In graduate school or teacher training courses, the literature review forms the foundation for much of the work, and the researcher is a student who is assigned a survey to get a grounding in SLA. If the researcher is conducting the review to support his or her own research project, there should be enough information to shape the research idea and focus it into a research question. As the project continues, the researcher will probably need to periodically refer back to the survey for support. It is also necessary to add to and edit the review as the research matures, the researcher's perspective changes, and the project focus narrows.

Summary

This paper endeavored to give the novice researcher some of the guidance and information needed to conduct a survey of SLA literature. The literature review is the most fundamental type of research. It forms the basis for much of the work at graduate school, and is essential for teacher training courses. It provides the necessary background to frame research projects and leads researchers towards their research goals. Performed early in a research project, it helps build the foundation on which to base further inquiry. As the research progresses, the literature review broadens the perspective of the study and helps to shape the research interest into a focused idea. The review requires the researcher to be a sleuth and a critic. It makes teachers look at the theoretical underpinnings of their profession and subsequently leads to greater understanding and professional development.

References

Baskin, R. (1994). Research in Japan. *The Language Teacher, 18* (11), 14,15,100.

Brown, J. D. (1988). *Understanding research in second language learning*. Cambridge: Cambridge University Press.

Brumfit, C. (1995). Teacher professionalism and research. In G. Cook and B. Seidlhofer (Eds.), *Principle and practice in applied linguistics: Studies in honor of H. G. Widdowson*, (pp. 27–41). Oxford: Oxford University Press.

Busch, M. (1994). Using Japanese libraries to do second language research. *The Language Teacher, 18* (2), 16–17.

Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.

Larsen-Freeman, D. and Long, M. (1991). *An introduction to second language acquisition research*. London: Longman.

Nunan, D. (1992). *Research methods in language learning*. Cambridge: Cambridge University Press.

Seliger, H. & Shohamy, E. (1989). *Second language research methods*. Oxford: Oxford University Press.

Chapter 5

Designing a Language Study

James Dean Brown

University of Hawaii at Manoa

This paper introduces some of the overarching issues in second language research. They are issues which must be addressed before conducting a study so that the researcher can avoid conceptual pitfalls that may cripple the study later on. The discussion will begin with the considerations involved in sampling a group, or groups, of subjects to be used in a study. Next, the different types of variables that researchers define in a study will be covered. Then, some of the research designs that can be used in second language studies will be explored. In addition, the factors which may jeopardize the internal and external validity of language studies are covered. Finally, the ethical issues involved in collecting data, conducting research, and reporting the results will be discussed.

Sampling

In language studies, it is often necessary to use sampling techniques. To understand why such techniques are necessary, it is first important to grasp the difference between a population and a sample. In research, a *population* can usually be defined as the entire group of language speakers or learners that the researcher wants to study. Unfortunately, few researchers have the resources to study, for example, the entire population of ESL students studying ESL in American universities, or the entire population of EFL students in the world, or even all of the male chemistry students from Germany who are studying in the United States. As a result, most researchers prefer to use a *sample*, that is, a subgroup of the students

representative of the given population. By using a sample, data can be practically and effectively collected, sorted, and organized. There are two basic strategies that are generally used in language studies for selecting samples from populations. These strategies are called random sampling and stratified random sampling. For both approaches, the purpose is to create an accurate sample, or subgroup, which can be said to be representative of the population as a whole.

Alternative Sampling Strategies

The underlying principle in ***random sampling*** is that each individual member of the population must have an equal chance of being selected into the sample. Three steps can be used to insure such equality of chance:

1. Clearly identify the population in which the researcher is interested.
2. Assign an identification number to each member of the population.
3. Choose the subjects for the sample on the basis of a table of random numbers.

A ***table of random numbers*** is a list of numbers, usually generated by a computer, that contains no systematic patterns. Most introductory statistics books contain such a list (for example, see Appendix A in Shavelson, 1981). Using a table of random numbers leaves the choices of who will be included in the sample up to a dispassionate and random table of numbers, rather than up to the researcher who may have subtle biases (conscious or unconscious) that could affect the results of the study. Once a large enough number of subjects is randomly selected, the resulting ***random sample*** can be assumed to be representative of the entire population from which they were drawn (Brown, 1988, pp. 111–113).

Other, more readily available techniques can be used to obtain a random sample. For example, the researcher might like to pull numbers out of a hat, use a deck of cards, or repeatedly throw a pair of dice in selecting subjects for a sample. Any technique wherein each member of the population has an equal chance of being selected, thereby ruling out biases on the part of the researcher, will be acceptable for random sampling, whether the sampling be for selecting subjects from a population for inclusion in a study, or for separating them into subgroups within the study itself.

Another strategy that is sometimes used in language studies is called *stratified random sampling*. In this case, four steps are usually used:

1. Clearly identify the population in which the researcher is interested.
2. Identify the salient characteristics of the population (called *strata*).
3. Randomly select members from each of the strata in the population (using a table of random numbers or other techniques described above).
4. Check to insure that the resulting sample has about the same proportions of each characteristic as the original population.

For instance, in the population of all ESL students studying at the University of Hawaii at Manoa (UHM), it might be useful to identify subgroups, or strata,

within the population based on the following characteristics: gender (male or female); country of origin; native language; academic status (graduate, undergraduate, or unclassified); and major (science, humanities, or undeclared). Given correct information about the proportions of these characteristics in the population of ESL students at UHM, the researcher could then randomly select from each of the strata in proportion to those population characteristics. The sample that results would intentionally take on the same proportional characteristics found in the entire population. Creating a stratified random sample still requires random sampling, but has the advantage of providing a certain degree of precision to the representativeness of the resulting sample—a fact which facilitates the use of the identified characteristics as variables in the study.

Decisions about which strategy (random or stratified random) to employ in a particular study must be reached rationally, and in advance. There are several considerations that must be kept in mind in making such decisions. First, it is generally useful to employ stratified random sampling when the population in question is fairly heterogeneous in nature. The concern is that random sampling might not provide for selection from each of the strata, or subgroups, in the population. Second, a stratified random sample becomes imperative when the samples involved will be small or the groupings within the study will be unequal in size. Third, it must be remembered that, if properly conducted, stratified random sampling has the advantage of letting the characteristics of the population determine which strata will be sampled. Hence, the stratified strategy is useful if the study will focus on the groups' characteristics as moderator or control variables (see Brown, 1988, pp. 11–18).

Alternatively, if the samples involved will be fairly large, straightforward random sampling can be employed. Random sampling is much easier to perform since there is no need to define the characteristics of the population. It is only necessary to assume that the sample represents the population from which it was taken. This assumption is widely accepted in research circles even though it is counter-intuitive for some language teaching professionals.

Sample Size

One of the first questions that will arise with regard to sampling is: How big must a sample be to be considered large enough? There is no easy answer to this question. However, it is clearly true that a large sample is better (in the sense of "more representative") than a small one. Consider a sample which includes all but 1% of a population of 1,000 language students (that is, a sample that contains 99% of the population). It is likely that such a sample is more representative of the population than one containing only 1% of it or 10% or 30%. However, knowing this does not answer the question of how big a sample must be to be considered large. Unfortunately, sample size decisions depend on the situation involved in the study as well as upon the types of statistics that will

be used. Statistics teachers will often give rules of thumb like the sample size should be at least 28 (or 30) per group or per variable. This is not bad advice *per se*, however, such rules of thumb are usually vague and imprecise, and in any case are conveying the minimum number that you will need for correctly applying many of the statistics that come up in research.

Another point of view is that, instead of estimating the bare minimum number of subjects, the researcher should be estimating the minimum number of subjects that would be necessary for a statistically significant result to be obtained (if it really exists in the population) given the application of a particular statistical procedure under the conditions of the study that is being planned. Such estimations can be made by using *power analysis*. One thing that power analysis can be used for is to analyze the relationship between the probability of finding a statistically significant result and the sample size given a particular set of expected results. If, for instance, a researcher wanted to estimate the number of subjects that would be necessary to find a statistically significant difference between the means of two groups of subjects, it could be done mathematically on the basis of pilot data, or other previous research that may be available in the literature. Such estimates can be made for a variety of the statistical procedures used for mean comparisons, correlation, and regression, as well as comparisons of frequencies (for more on power analysis, see Cohen, 1988; Kraemer & Thiemann, 1987; Lipsey, 1990). Unfortunately, power analysis is mathematically complex. However, there is computer software available (e.g., Borenstein & Cohen, 1988) that can resolve this problem.

In short, when thinking about sample size, the best strategy is to make sure that the population is clearly defined, and that the sampling procedures make sense. If pilot data or other previous research is available, it will prove helpful to use power analysis to estimate the sample size that is necessary to find a significant effect if it exists. If pilot data are not available, you may have to design your study such that the samples involved "seem" large enough to be representative, while keeping in mind that a good rule of thumb is the larger the sample size the better. The issues involved in sampling are somewhat subjective, and must in part be left up to the researcher. Sampling procedures are important partly because of the way that they affect the generalizability of the study.

The *generalizability* of a study can be defined as the degree to which the results are meaningful beyond the study itself with regard to the entire population in question. If the sampling techniques have been properly conducted and the sample is large enough, there should be no question in the researcher's mind (or in a reader's mind) as to the degree to which the sample represents the population. If there is some question, then the sampling techniques should be improved or the sample sizes increased, or both. (For more information on sampling and its use in language studies see Brown, 1988; Hatch & Lazaraton, 1991.)

Different Types of Variables

A *variable* is anything that can vary in a study. However, research is largely the study of what happens when variables are systematically manipulated in planned combinations. There are essentially five roles that variables can play in a study: dependent variables, independent variables, moderator variables, control variables, and intervening variables.

The *dependent variable* in a study is the variable of primary focus. It can also be thought of as the variable that is measured and studied to determine if other variables have an effect on it, or are related to it. The *independent variable* in a study is the variable that has been selected by the researcher in order to study its effect on the dependent variable (hence, the independent variable is sometimes also called the manipulated variable). For instance, for a research question like "What is the effect of X on Y?", X is the independent variable and Y is the dependent variable. Or, for a research question like "How well does X predict Y?", X is the independent variable and Y is the dependent variable.

The relationship between the independent and dependent variables is central to any study. However, sometimes the researcher will also want to include a *moderator variable* in order to determine the effect of the moderator variable on the relationship between the dependent and independent variables. Thus, if a moderator variable were included, a question like the following could be posed: "What is the effect of X on Y when Z is present or absent?" In this last case, X is the independent variable, Y is the dependent variable, and Z is a moderator variable.

In language research, there are usually variables other than the dependent, independent, and moderator variables which cannot be included in the design or otherwise directly studied. Nonetheless, these variables must be accounted for often as control variables. *Control variables* are variables which are eliminated from the study, held constant, or otherwise kept from interfering with the study of the central relationship between the independent and dependent variables. For instance, in a study of the effect of Method A on English language proficiency (as measured by TOEFL), the researcher might compare the TOEFL scores of two groups, one who had been taught by Method A and another who had received no instruction, a control group. The researcher would be most interested in the relationship between the independent variable, Method, and the dependent variable, English Language Proficiency. However, there are a number of variables which might interfere with the relationship between Method and English Language Proficiency: gender, intelligence, aptitude, years of language study, etc. The researcher might choose to control gender by eliminating all males from the study. The researcher might further choose to use only students who had studied six years of English to hold the years-of-study variable constant. Random selection can also be used to create groups that are theoretically equal on all variables except those being manipulated as independent, dependent, and moderator variables.

Perhaps the most confusion is caused by the term "intervening" variable because it is used in two distinctly different ways. On the one hand, *intervening variable* is used to describe the construct which underlies the relationship between the independent and dependent variables. For instance, in the example study above on the effect of Method A on English Language Proficiency, the researcher might label the construct underlying the effect as "method effect" or "learning" or "language acquisition" depending on how it is conceptualized.

On the other hand, *intervening variable* is used to describe a variable that is unanticipated in a study, yet surfaces as being a possible explanation for the relationship between the independent and dependent variable. In the example study, it might turn out that any difference discovered in the proficiency scores of the Method A and Method B groups were caused by an unanticipated intervening variable rather than by the methods themselves. For instance, it might turn out that the teacher of the Method A class was just a better teacher than the teacher of the control group. Thus, a teacher effect turns out to be a possible intervening variable in the sense that it was unanticipated yet has potential explanatory power.

Research Designs

To understand the basic designs that are used in quantitative language studies, it will first be necessary to define some of the fundamental terms that are used. The first idea that must be understood is that of a *treatment*. A treatment is something that the experimenter does to one group so as to study the effects of the treatment on the people involved. A treatment may be a specific teaching strategy, application of a set of materials, use of a particular reward system, or any other experience that the researcher wants to apply to the subjects for the study. Typically, for the sake of comparison, one group receives the treatment while another group does not. Thus, the subjects are divided into two or more groups: a control group and one or more experimental groups. The *control group* usually receives no treatment, or a placebo (some substitute that is predicted to have no effect), while the *experimental group* receives the treatment. In a language program, the treatment is likely to be some aspect of the language teaching or learning experience.

The reason for administering a treatment to the experimental group and nothing to the control group is to determine whether the treatment has had an effect. In order to do so, one or more observations must occur which allow for comparisons of the two types of groups. These *observations* may take many forms. In quantitative studies, observations may be simple tallies, rankings, or test scores. The point in making observations is that something of interest to the researcher must be observed or measured so that comparisons can be made between the control and experimental groups. Naturally, whatever is observed

or measured must be related to the treatment. Thus, in a language program, if the treatment was some form of pedagogy, you might be interested in observing the language achievement test scores in order to determine the effect of the treatment on achievement.

It is important to note that studies involving anything other than examination and description of test scores are difficult to conduct. The study may be designed in an airtight manner (difficult in any teaching or learning situation), but in addition, considerable knowledge of statistics must be applied—usually more than the knowledge provided in one or two statistics courses. This warning is meant to encourage budding researchers to seek adequate guidance in designing quantitative studies and analyzing the statistical results.

The two sections that follow will explain two categories of quantitative studies: true experimental designs and quasi-experimental ones. This is a very useful distinction explained much more fully in Campbell and Stanley (1963).

True experimental designs are the most controlled language studies. They must be carefully planned from beginning to end. Hence, they are the closest thing in language studies to what most teachers believe scientific experiments are like. One of the keys to identifying a true experimental design is that the subjects in the study must be randomly selected from the population being studied, and randomly assigned to the treatment or control group. *Randomly* is used here strictly in the sense that it was defined above. As described in the earlier section on *Sampling*, this must be done so that every member of the population has an equal chance of being selected. If these procedures are followed and the resulting groups are large enough, the researcher is justified in assuming that the two groups have very much the same characteristics. Thus, true experimental designs have random selection as a precondition. The same thing is true for posttest only designs, pretest-posttest designs, or any combination of the two.

The *posttest-only design* (one type of true experimental design) is particularly dependent on random selection because it is assumed on the basis of sampling theory that the experimental and control groups are equivalent at the outset of the study. Such a study is designed as shown in Figure 1. Notice that step one is to use random selection to create equivalent groups. The experimental group receives the treatment, while the control group does not (or receives a placebo). Both groups are then observed on the same scale and the performances of the two groups are compared. If the experimental group has significantly higher performance than the control group, arguments can then be built that the treatment has an effect. The degree to which such claims can be made will naturally depend on the magnitude of the differences in performance.

The *pretest-posttest design*, while it also assumes random selection of the two groups, allows the researcher to check the equivalence of the two groups at the beginning of the study, usually a pretest of some sort. Such a study would typically be laid out as shown in Figure 2. This additional step allows for checking the

Figure 1
True Experimental Design, Posttest Only

Step 1		Step 2		Step 3	Step 4
Random Experimental Group	→	Receives Treatment	→	Observation	Performances Compared
Random Control Group	→	Receives No Treatment	→	Observation	

equivalence of the two groups in Step 2, but also allows for studying the amount of gain that has been made by each group between Steps 2 and 4. This potential for studying gain allows the researcher to consider additional issues. For example, if there is a difference between the two groups on the posttest, the researcher can study whether the difference is as large as the difference between the pretest and posttest performances of the experimental group. If this is not true, the observed differences may have some source other than, or additional to, the treatment. Thus, the pretest-posttest design is generally more powerful than the posttest only design because more inferences can be drawn. Pretest-posttest designs can become much more complex including various types of treatments used simultaneously and various observation techniques used in the same study.

Figure 2
True Experimental Design, Pretest and Posttest

Step 1		Step 2		Step 3		Step 4	Step 5
Random Experimental Group	→	Observation (Pretest)	→	Receives Treatment	→	Observation (Posttest)	All Observations Analyzed
Random Control Group	→	Observation (Pretest)	→	Receives No Treatment	→	Observation (Posttest)	

From a practical point of view, true experimental designs are often doomed in real language teaching settings. First, students are rarely randomly selected. Thus, many researchers are working either with what is called an *intact group* or with the entire population of students when they set out to do a study. Second, the researcher cannot set aside half of the students, randomly selected or otherwise, to receive no language training, or a placebo. Either students want the training or they do not, and language researchers are seldom in the moral or monetary position to simply withhold treatment (training) from one half while the other half receives training. As a result, language researchers are more likely to turn to what is called a quasi-experimental design. *Quasi-experimental designs*, though less than perfectly controlled, provide useful alternatives to true experimental designs. Quasi-experimental designs are adequate for the purposes of studying many language issues—particularly if no sweeping claims are going to be generalized from the results. According to Campbell and Stanley (1963), the main characteristic that makes a quasi-experimental design more practical for language studies is that the researcher has more control over the collection of data in terms of scheduling and deciding who will participate. However, the designs are weaker, and the results must be interpreted very carefully. Three types of quasi-experimental designs will be presented here: pretest-posttest designs (without control group), time series designs (without control group), and nonequivalent groups designs.

The *pretest-posttest design without control group* is like the pretest-posttest design discussed above except that it lacks a control group. Such a design is shown in Figure 5. This type of design could be used as follows: A general proficiency pretest could be given at the beginning of a language program (the treatment) and again as a posttest at the end of the program. If there is a large gain in average scores between the beginning and end of the program, it might be judged as a success. However, because there is no control group in such a study, the researcher can never know for sure that the gain was not a result of language exposure outside of the program, or a result of a testing effect (that is, the effect of having taken the test twice), or a result of some other undetermined factor. In other words, the observed gains in scores may have been due to factors other than the learning that took place in the program.

Figure 5
Quasi-Experimental Design, Pretest and Posttest

Step 1	Step 2	Step 3	Step 4	Step 5
Experimental Group	→ Observation (Pretest)	→ Receives Treatment	→ Observation (Posttest)	] Observations Compared

Time series designs are more elaborate versions of the pretest-posttest design. The only striking difference is that, in lieu of one pretest and one posttest, a series of observations, or tests, are made. Then, a treatment is inserted in the middle of this series. Such a design is described in Figure 4 (in which "O" stands for Observation). In a time series design, the researcher can claim that the potential consequences of the testing effect mentioned above are controlled in that all students are made thoroughly familiar with the format and content types on the observation instruments long before the treatment comes into the picture. One problem that arises with this type of design is that it sometimes calls for the development of numerous instruments, all of which must be very similar in what they measure.

Figure 4
Quasi-Experimental Design, Time Series

O 1	→ O 2	→ O 3	→ O 4	→ Treatment	O 5	→ O 6	→ O 7	→ O 8		All Observations Analyzed

The *nonequivalent groups design* is different from the true experimental pretest-posttest design only in that the subjects are not randomly selected into the experimental and control groups. Such a design is shown in Figure 5. Because the groups are not randomly selected, they cannot be assumed to be equivalent at the beginning of the study. As a result, the equivalence of the groups must be checked in Step 2 (or otherwise controlled statistically). If it is possible to set up a control group in this manner and the groups do indeed prove to be equivalent at the beginning of the study, the nonequivalent groups design can prove fairly powerful. However, if such a control group cannot be established, the quasi-experimental version of the pretest-posttest design (Figure 3) may be the most effective design that can be used.

Figure 5
Quasi-Experimental Design, Nonequivalent Groups

Step 1	Step 2	Step 3	Step 4	Step 5
Experimental Group →	Observation (Pretest) →	Receives Treatment →	Observation (Posttest)	All Observations Analyzed
Control Group →	Observation (Pretest) →	Receives No Treatment →	Observation (Posttest)	

There are many other types of complex designs (see Campbell & Stanley, 1963, or Tuckman, 1978), and numerous ways of grouping and analyzing the results of those designs (see for instance, Keppel, 1973; Kirk, 1968; Pedhazur, 1982; Tabachnick & Fidell, 1989).

Validity

The validity of a study can be defined as the degree to which the results can be accurately interpreted and effectively generalized. The first part of this definition—the degree to which the results can be accurately interpreted—is often referred to as *internal validity*. The second part, the degree to which the results can be generalized, is often labeled *external validity*. Table 1 lists the different factors that can affect the validity of a study (after Campbell & Stanley, 1963).

Internal Validity

The eight threats to internal validity, listed above, are variables that must be controlled in designing a study so that the results can be accurately interpreted.

History includes anything that happens to the subjects, other than the intended treatment, between the observations in a study. For example, for the design shown in Table 1, history would be anything, other than the treatment,

Table 1
Factors Threatening Validity (after Cambell & Stanley, 1963)

Type of Validity Factor
Internal Validity
1. History
2. Maturation
3. Testing
4. Instrumentation
5. Statistical regression
6. Selection bias
7. Experimental mortality
8. Selection-maturation interaction
External Validity
9. Reactive effects of testing
10. Interaction of selection biases and the treatment
11. Reactive effects of experimental arrangements
12. Multiple treatment interference

Ci

that occurs between the pretest and posttest for either the experimental or control group in the True Experimental Design, Pretest, and Posttest.

Maturation refers to any of the processes in the subjects' lives that occur because of the passage of time and might interfere with interpretation of the results of a study. For instance, fatigue, hunger, aging, changing schools, or passage through puberty would all be maturation factors that the researcher might want to consider.

Testing describes any influence that taking one test has on the scores of another test. For instance, taking the pretest shown in Table 1 might affect the scores on the posttest. The testing effect might be particularly pronounced if the type of test involved were completely new to the subjects involved. Consider a group of subjects who had never taken a cloze test before. If one were administered as a pretest, the subjects might learn test taking strategies that would make them more comfortable and make them score higher on a subsequent posttest, regardless of any treatment that was administered.

Instrumentation involves the impact of variations in the tools of measurement or problems with the reliability of those tools (for much more on this latter topic, see Brown, 1995a, 1995b) on the obtained measurements, or scores. For example, a problem of instrumentation would arise if version A of a test was used in the pretest, but version B was used on the posttest. The problem is that any differences in performance could be due to discrepancies in the versions of the test (the instruments) rather than to any treatment involved.

Statistical regression describes the moderating effects of selecting groups with extreme scores, either very high, very low, or both. Under such conditions, the probability is that students with high scores will tend to score lower (i.e., closer to the average score), while students with very low scores will tend to score higher (i.e., closer to the average score) for reasons having nothing to do with any treatments involved.

Selection bias describes the impact of selecting subjects into the groups of a study for reasons other than chance. For instance, if the subjects for the experimental group in Table 1 were selected from students in 8:00 a.m. ESL classes, while the subjects in the control group were selected from students in 4:00 p.m. ESL classes, there might be differences in the groups based on class time preference that have nothing to do with the treatment involved.

Experimental mortality refers to the influence of subjects dropping out of one or more of the groups in a study. For example, in Table 1, some subjects in the control might be present for the pretest but absent for the posttest. These absences might cause differences in the results that had nothing to do with the treatment.

Selection-maturation interaction describes the effect of the maturation and selection bias variables (defined above) acting together.

External Validity

The four threats to external validity listed above can affect the generalizability of the results.

Reactive effects of testing and treatment describe the influence of taking a pretest on the sensitivity of the subjects to the treatment. For instance, if the treatment involved the use of cloze tests to practice reading prediction and the pretest was also a cloze test, the pretest might affect the subjects' sensitivity to the treatment. In other words, the generalizability of the results might be in question because the results depend on the use of a particular test.

Interaction of selection biases and the treatment. If there is some relationship between the group from which the subjects were selected and the effects of the treatment, interactions are said to exist between selection biases and the treatment variable. In other words, in any study, there is the possibility that any effects that are found are only true for the population from which the groups were selected. It is also possible that the characteristics of that particular population may cause the treatment to be effective where it would not be in another population. In such a situation the selection bias would be interacting with the treatment and thus affecting the generalizability of the results.

Reactive effects of experimental arrangements. This refers to the impact of the fact that the treatment was applied under experimental conditions rather than real world conditions. For example, some pedagogical techniques might appear to work very well as a treatment under classroom conditions but have no correspondingly beneficial effect on the students' use of the language in the real world. The generalizability of the results to the real world would be in question.

Multiple treatment interference refers to the effects of applying more than one treatment to the same subjects. Under these conditions, the effects of one treatment cannot be disentangled from the effects of others, and thus the results cannot be generalized to situations that do not contain the multiple treatments.

Multiple Threats to Validity

Unfortunately, threats to validity in research are seldom as simple as Table 1 would suggest. This is because there may be numerous threats to validity operating at the same time. Since the overall confusion caused by simultaneously having multiple threats to the validity of a study may well be synergistic, it is crucial that researchers guard against and control any and all of these problem factors, preferably while planning a study. (For more information on factors that threaten the validity of a study and how to control them, see Brown, 1988; Campbell & Stanley, 1963; Haich & Lazaraton, 1991; Tuckman, 1978.)

Ethics

Ethics in social science research have been considered from a number of perspectives. For an overview of this work see Kimmel (1988). Over the years, various organizations associated with social sciences research have provided guidelines for their memberships. For example, the American Psychological

BEST COPY AVAILABLE

Association has provided various kinds of guidelines for the ethical conduct of research (American Psychological Association, 1953, 1981, 1982, & 1985). According to Kimmel (1988), ethical problems in social sciences research may have a number of the following characteristics:

1. The complexity of a single research problem can give rise to multiple questions of proper behavior.
2. Sensitivity to ethical issues is necessary but not sufficient for solving them.
3. Ethical problems are the results of conflicting values.
4. Ethical problems can relate to both the subject matter of the research and the conduct of the research.
5. An adequate understanding of an ethical problem sometimes requires a broad perspective based on the consequences of research.
6. Ethical problems involve both personal and professional elements.
7. Ethical problems can pertain to science (as a body of knowledge) and to research (conducted in such a way as to protect the rights of society and research participants).
8. Judgments about proper conduct lie on a continuum ranging from the clearly unethical to the clearly ethical.
9. An ethical problem can be encountered as a result of a decision to conduct a particular study or a decision not to conduct the study.

Commandments

In language related research, some of the most important ethical and professional issues might best be summed up by ten straightforward commandments (adapted from Brown 1984). These commandments cover the researcher's ethical and professional responsibilities with regard to the participants, analyses, and audience of a study.

Participants

I. Thou shalt not abuse thy subjects in any manner including abuses of their persons, time, or effort, and thou shalt obtain thy subjects' informed consent if required by thy institution.

II. Thou shalt not abuse thy colleagues by collecting data from their students without permission, or by using too much precious class time.

III. Thou shalt reward thy subjects' and colleagues' efforts at least by giving them feedback or information on what happened in the study.

Analyses

IV. Thou shalt guard against consciously or subconsciously modifying thy data so that the results support thy views and prejudices.

V. Thou shalt select the appropriate statistical tests.

VI. Thou shalt check the assumptions that underlie all statistical tests.

C4

Audience

VII. Thou shalt explain thy research clearly so that it can be understood by thy readers.

VIII. Thou shalt organize thy report using conventional sections, headings, and other conventions (see American Psychological Association, 1994) so that thy readers can easily follow thy study.

IX. Thou shalt interpret thy results carefully guarding against the temptation to over-interpret, or generalize beyond that which thy results warrant.

Above All Else

X. Thou shalt continue to learn, read, and grow as a researcher so that thou can better serve thy field.

Stating the ethical issues in the form of commandments may at first seem to be intended as tongue-in-cheek humor, but these are not to be taken lightly. Indeed, the entire enterprise of research in language studies hinges on cooperation between subjects, colleagues, researchers, and readers. Researchers should avoid contributing to the already abundant negative feelings about statistical research.

Conclusion

This paper began with a discussion of the issues involved in sampling a group, or groups, of subjects to be used in a study. Then, data collection instruments were examined in terms of the four scales of measurement that can be used. Next, a number of research designs were explored. Then, the factors which may jeopardize the internal and external validity of language studies were surveyed. Finally, the ethical issues involved in collecting data, conducting research, and reporting the results were covered. In short, a great many crucial issues have been covered in this paper—issues that must be thought through before conducting a study. A little effort spent in the planning stages of a study can save the enormous amount of energy necessary to recover if things begin to come unraveled after the study has begun.

References

American Psychological Association. (1953). *Ethical standards of psychologists*. Washington, DC: American Psychological Association.

American Psychological Association. (1981). *Ethical principles of psychologists*. Washington, DC: American Psychological Association.

American Psychological Association. (1982). *Ethical principles in the conduct of research with human participants*. Washington, DC: American Psychological Association.

American Psychological Association. (1994). *Publication manual of the American Psychological Association* (4th ed.). Washington, DC: American Psychological Association.

American Psychological Association. (1985). *Standards for educational and psychological testing* (3rd ed.). Washington, DC: American Psychological Association.

Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.

Borenstein, M., & Cohen, J. (1988). *Statistical power analysis: A computer program*. Hillsdale, NJ: Lawrence Erlbaum.

Brown, J. D. (1983). An exploration of morpheme-group interactions. In K. M. Bailey, H. Long, & S. Peck (Eds.). *Second language acquisition studies*. Cambridge, MA: Newbury House.

Brown, J. D. (1984). Moral, ethical and social considerations in quantitative TESOL research. Paper presented at the TESOL Convention, Houston, TX.

Brown, J. D. (1988). *Understanding research in second language learning: A teacher's guide to statistics and research design*. London: Cambridge University.

Brown, J. D. (1995a). *The elements of language curriculum: A systematic approach to program development*. Boston, MA: Heinle & Heinle.

Brown, J. D. (1995b). *Testing in language programs*. Englewood Cliffs, NJ: Prentice Hall.

Brown, J. D., Knowles, M., Murray, D., Neu, J., & Violand-Hainer, E. (1992). *A survey of research and practice in the TESOL organisation*. Paper presented at the TESOL Convention, Vancouver, Canada.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Hatch, E. & Lazaraton, A. (1991). *The research manual: Design and statistics for applied linguistics*. New York: Newbury House.

Keppel, G. (1973). *Design and analysis: A researcher's handbook* Englewood Cliffs, NJ: Prentice Hall.

Kimmel, A. J. (1988). *Ethics and values in applied social research*. Newbury Park, CA: Sage.

Kirk, R. E. (1968). *Experimental design: Procedures for the behavioral sciences*. Belmont, CA: Brooks/Cole.

Kraemer, H. C. & Thiemann, S. (1987). *How many subjects?* Newbury Park, CA: Sage.

Krashen, S. (1977). Some issues relating to the monitor model. On *TESOL '77*. Selected papers from the 11th annual TESOL convention, Miami (pp. 144–158). Washington, DC: TESOL.

Lipsey, M. W. (1990). *Design sensitivity: Statistical power for experimental research*. Newbury Park, CA: Sage.

Pedhazur, E. J. (1982). *Multiple regression in behavioural research: Explanation and prediction* (2nd ed.). New York: Holt, Rinehart and Winston.

Shavelson, R. (1981). *Statistical reasoning for the behavioural sciences*. Boston: Allyn and Bacon.

Tabachnick, B. G. & Fidell, L. S. (1989). *Using multivariate statistics* (2nd ed.). New York: Harper Collins.

Tuckman, B. W. (1978). *Conducting educational research*. New York: Harcourt Brace Jovanovich.

Chapter 6

Methods for the "Research Challenged"

Christopher Jon Poel

Musashi Institute of Technology

Robert M. Homan

International Christian University

Research, simply put, is the gathering of information (data) that will help to answer a question about the world around us. It can be broken down into two categories, secondary and primary. Primary research "is derived from primary sources of information (e.g., a group of students who are learning a language), rather than from secondary sources (e.g., books about students who are learning a language)" (Brown, 1988, p. 1). Research also "involves planned and systematic inquiry" (Seliger & Shohamy, 1989, p. 12), which differentiates it from common sense.

This is the same explanation we see in all the research manuals, and it is not what most *teachers* want to hear. Many language teachers have an ingrained fear of the very concept of research. They may say, "Research is for people with PhDs in the field. I just want to teach my courses." Others may say, "Research requires statistics, and I don't want to have anything to do with statistics."

However, research, whether we are willing to acknowledge it or not, is present in all facets of our teaching experience. How many times have you tried a new way of teaching vocabulary? How did you evaluate the efficacy of the technique? Did you compare test scores? Did you ask a few questions of the

students to see how they liked it? Did you overhear some students' remarks while they were doing the lesson? Did you look at their faces and notice that their eyes were lighting up with discovery—or clouding over with confusion? If you have done any of the above, you have started down the road toward doing research. Of course, you did not formalize your hypotheses, write up a proposal, and submit it for publication, but you did a form of research all the same.

Have you ever had to teach a pronunciation class and wondered which sounds to teach the students? Well, in order to find out, you probably did a kind of needs analysis. In other words, you may have interviewed some of the students to get a general feel for what types of sounds give them major difficulties. You may have given them a quiz in which they were to mark the correct answer to a question about how to make a particular sound. You may have given them a listening test in which they were to determine which word in a minimal pair was spoken. Or, you may simply have paid careful attention to their pronunciation while doing another lesson. No matter how you went about it, the fact is, the needs analysis you did is a form of research.

What we intend to do in this paper is to take some of the mystery out of the word "research" by showing you some ways you can carry it out without radically changing what you normally do. In other words, you can do a lot of useful and interesting studies without ever leaving your classroom.

Our goal is that this paper will help you to realize that research is something you *can* do, because you have been on the brink of doing it all along. The paper is divided into sections, with each covering a particular methodology. The sections will address the questions: What is it? When do I do it? How do I set it up? How do I analyze the results? And, where do I go from here?

Action Research

What is it?

Action research (AR) could be described as classroom research on a shoestring. It is a kind of research that is accessible to every teacher. Simply put, to do action research all you need is a classroom, a problem, and an idea or two on how to solve the problem. As a good teacher that is exactly what you do every day. But what sets AR apart is that it is done in a more systematic manner. Action research formalizes what we do unconsciously every time we step into the classroom.

When do I do it?

Let us say that you are having trouble with a class; the students are not responding to your well-prepared lessons, and you know the lessons are good because all your other classes are doing well. Or, a colleague tells you about a new way of teaching an old lesson, and you decide to try it. Or you are using a new textbook this year, one that you had no say in choosing. Is the book any good?

And how can you make it better? These are some of the more obvious times when action research can be a useful tool.

Whereas many types of research—most notably experimental research with all of its statistical analyses—are usually carried out by experts, action research is essentially the teacher's ball game. In cooperation with other teachers and with students—and even with full-time researchers—a problem is identified and worked on. From this perspective, the best time to do AR is any time. That is what makes this type of research so attractive to the classroom teacher. Talk to your colleagues, decide what to teach, try it, evaluate it, revise it, try it again, evaluate it again. Hey, you have just done your first action research study.

We would like to point out the usefulness of a video camera. Even though it is troublesome to bring a camera into the classroom, set it up, and get usable video footage, the positives definitely outweigh the negatives. You do not have to watch the tape and transcribe all the utterances and actions; simply look at it. You will discover a lot of interesting things you never thought of, because you will be looking at it from a different perspective. It is also easier to consult with your colleagues, because you can show them the relevant parts of the lesson as they happened, as opposed to trying to explain how things went.

How do I set it up?

One aspect of AR that is often ignored is the need for collaboration. It cannot happen when teachers jealously guard their secrets for fear that they will be stolen. Get together with your fellow teachers, outline your problem, brainstorm some solutions, choose the most likely one, and plan your study.

Once you have identified your problem and potential solution, it is time to conduct your research. As a guideline for what you should look for when observing your students, Knowles (1990, pp. 7-10) offers the following questions:

- What did the students actually do?
- What were they learning?
- How worthwhile was it?
- What did I do?
- What did I learn?
- What do I intend to do now?
- What would I recommend to my colleagues?

In essence, these questions replace the formal hypotheses and research questions found in experimental research. However, that does not mean that action research can be sloppy. Thinking about these questions combined with careful observation of what happens in your classroom is essential for carrying out a successful AR study. So, do not just pass over these questions when you do some AR—write them down along with the answers. It is just like the essay outlines you have your students write—not much fun, but really useful.

How do I analyze the results?
As with other aspects of action research, evaluation is also quite informal. Observe what happens during the lesson, share your observations with your colleagues, and discuss how effective your solution was and how to refine it. If the problem has disappeared, great! If not, go back to the drawing board, and consider other possible solutions.

Where do I go from here?
Action research is a great starting point for teachers who are allergic to research. It does not take a lot of resources or extra time, just a little forethought and collaboration with your fellow teachers. So, read the other papers in this volume, talk to your colleagues, observe your classes, and you will be on your way. And if you get bitten by the research bug, read on to find out how to dive into the ocean of hard research head first.

Survey Research

What is it?
A second type of research that teachers often do unconsciously is survey research. Have you ever asked students for their opinions about a lesson? Have you asked them what they feel they need to study, what their weak points are, or what they want to discuss next week? If so, you have done survey research. Many teachers may not consider it so, but if research is the gathering of data to answer specific questions, then surveys certainly fit the bill.

Survey studies "focus on a group's attitudes, opinions, and/or characteristics" (Brown, 1988, p. 3) and are often exploratory in nature. The researcher's aim is to gather as much information as possible on a question rather than to confirm an answer. There are two basic types of survey research. The first is the use of questionnaires, which are especially useful in gathering data from a large number of people. The second type of survey is the interview, which is very useful in gathering detailed information from a smaller number of subjects. It is often used as a follow up to a questionnaire, but can also stand on its own.

The best way to decide if a survey study is right for you is to look at what you want to achieve; examine all the parameters, such as who you want to gather information from and how easily you want to be able to analyze the data once it is collected.

When do I do it?
Survey research is especially useful in collecting descriptive data. Suppose you want to collect data on the needs of your students, their opinions on some aspect of the lessons or class, ideas for future lessons, personal information, or anything of interest to you that may not be easy to derive from a test—then surveys are a likely candidate.

Questionnaires can be useful when you need to collect information from the students in your classes. As an example, consider a curriculum revision project undertaken at a vocational school. As part of the overall curriculum revision project, a needs analysis was developed in the form of a questionnaire, both for the students and for their future employers. The information from the questionnaire was then used to begin revising the curriculum. Once again, the information collected did not directly reveal the best curriculum but did provide much useful data for making informed decisions.

On the downside, if you do not have a captive audience (like an EFL class) the response rate may be low. If the percentage of people who respond is too low, the value of the study will be called into question. For instance, imagine that you send a class evaluation questionnaire to your students after the class has finished for the year—after all, you do not want the students to think that their responses are going to influence their final grades. After a month, you look at the responses that you received, and find out that 94% of the respondents thought your class was great, 94% said your teaching style was wonderful, and 94% believed your textbook was perfect. Then you notice that out of 150 students, only 17 responded, and 16 (or 94%) of those 17 were A students. Well, most likely you already know that the A students liked you. What you need to know is why the B, C, D, or F students did not respond to your lessons. An interview may be useful here because it allows you to focus on specific information. The one student who did not agree with the majority might be able to provide more valuable information than the other 16 students did. As such, it would behoove you to follow up the questionnaire with an interview, either by telephone or in person, with the student who may not have liked you, might not have gotten much out of the textbook, or could have benefited from a different style of teaching.

How do I set it up?

Setting up a survey is both easy and difficult. It is easy if you think that all you have to do is write up a bunch of questions, copy them, give them to your subjects, collect the papers, and read them. However, you need to consider the following: What information do you want to gather? How structured do you want it? Do you want open-ended questions so the subjects can write their answers freely, or do you want to limit the range of responses to make analysis easier? It is the actual writing of the questions that makes survey research difficult.

In this vein, Ary, Jacobs, and Razavieh, (1990, p. 422–424) have offered the following guidelines to help construct an effective questionnaire:

- Construct the instrument in such a way that it reflects quality. [The questionnaire should be carefully written, proofread, and laid out in order to ensure a high response rate.]
- Keep the questionnaire as brief as possible, so that it requires a minimum of the respondents' time.

- Make sure that the respondents have the information necessary to answer the questions.
- Phrase questionnaire items so that they can be understood by every respondent.
- Keep individual questionnaire items as short and simple as possible.
- Phrase questionnaire items so as to elicit unambiguous answers.
- Phrase questionnaire items so as to avoid bias that might predetermine a respondent's answer.
- Avoid questionnaire items that might mislead because of unstated assumptions.
- Make sure that the alternatives to each questionnaire item are exhaustive, that is, express all the possible alternatives on the issue.
- Avoid questions that might elicit reactions of embarrassment, suspicion, or hostility in the respondent.
- Avoid "double-barreled" questions that attempt to ask two questions in one.

To illustrate what this all means, consider the following examples of good and bad questionnaire items (Nunan, 1992, p. 143–144):

- Good: What do you think about the proposal that foreign languages should be compulsory in high school?
- Bad: Do you think that the concept of learner centeredness is utopian and unrealistic? [This is a leading question which attempts to predetermine the respondent's answer.]
- Good: Rank the following from 1 to 4 in order of preference. "I like to learn best by studying":
  - with the whole class
  - in small groups
  - in pairs
  - independently
- Bad: Would you prefer a short, non-award course (3, 4, or 5 sessions) with part-day release (e.g., Wednesday afternoons) and one evening per week attendance with financial reimbursement for travel, or a longer, non-award course (6, 7, or 8 sessions) with full-day release, or the whole course designed on part-time release without evening attendance? [Besides being extremely complex and confusing, this question attempts to gather information on more than one topic.]

Types of questions include open-ended (completion or fill-in questions where respondents write their answers freely), fixed choice (similar to multiple-choice test items), checklists ("Check all that apply"), ranking ("Rank the following from 1 to 5"), or scaled items ("On a scale of one to a hundred, how would you rate your teacher's choice of necktie?").

The scaled items are often written in the form of a Likert scale. Everyone who reads a popular magazine has probably run into this type of scale. For example:

How do you feel when a dentist is looking at you with drill in hand and gleam in eye?

1	2	3	4	5
uncomfortable				ecstatic

Likert scales are especially useful when you need information on the opinions or attitudes of the respondents on issues that can be dichotomized, that is, separated into two opposite attitudes. Some of the common dichotomies found in language research are very serious/slight, important/unimportant, like/dislike, and agree/disagree (Brown, ms), not to mention the ever-popular uncomfortable/ecstatic distinction, as illustrated above.

How do I analyze the results?

Survey data, in general, will result in descriptive analyses rather than statistical analysis. If your questions are open-ended, you could theoretically end up with an infinite number of different responses—assuming, of course, that you had an infinite number of subjects who responded. You would then use these responses to tease out any patterns of interest for further investigation, such as interviews for more detailed data or an experimental study. Fixed choice and Likert scale questions, on the other hand, would most likely be analyzed as percentages, as in this example:

Does your teacher arrive in the classroom on time?

1 = 75%	2 = 22%	3 = 3%	4 = 0%	5 = 0%
almost never				almost always

Again, one of the purposes in calculating the percentages is to discern patterns which merit further investigation. In our example, it would certainly be expedient to the researcher—or at the very least, the department chair—to investigate the situation with the late teacher.

Although this is the most common use for the data from a survey study, it is also possible to construct questionnaires or structure interviews to investigate the relationship between different variables, such as the teacher being late and the students' overall liking of the teacher. In the example, the students responded that they liked the teacher very much, most likely because their lesson was first period and they knew they could get away with being 30 minutes late for class!

Where do I go from here?

We suggest that you take a look at your courses and ask yourself if the students seem to like or dislike a certain aspect, for example a reading or a video that you use. Construct a questionnaire and evaluate the data. You may find that they rate it highly, in which case you would want to introduce more of the

same. Or, you may find that they rate it very poorly, and then you have to decide if the benefits of the material or activity offset the negative impressions of the students. At least you will have some sort of a starting point to work from.

Correlational Research

What is it?

According to Brown (1988), correlational studies "are designed to investigate the nature and strengths of functional relationships among the variables of interest to the researcher" (p. 126). Another definition was offered by Hatch and Farhady (1982, p. 192): "In correlational studies, researchers are interested in determining the degree of relationship between pairs of two or more variables. . . . In other words, correlational studies allow us to determine the extent to which scores on one test are associated with scores on another test." Correlational studies are pretty easy to comprehend, but the calculation of the "correlation coefficient" may give some people, including us, fits. Just keep in mind, though, that the calculations are the reason we pay big bucks for computers—and even bigger bucks to someone to run them for us.

Basically, what you want to do is see if two scores or measurements are related to each other and how strong the relationship is. For example, one may compare the scores of a vocabulary test with those of a spelling test to see if they are related. The characteristics, in this case vocabulary ability and spelling ability (as measured by the scores on the two tests), are well defined, and the scores certainly vary from student to student. Once the strength of the relationship is established, one could make predictions based on that particular relationship. It should be noted, however, that this is only a prediction and is not to be taken as something chiseled in stone. The problem with correlational studies is that they may not take unknown variables into account, or the variables may not be as well-defined as we imagine. Also, it bears remembering that a very strong correlation does not mean anything more than that there is a strong relationship—causation is not established through a correlational study.

When do I do it?

A correlational study can be set up almost any time, as long as you have two or more variables that you wish to investigate. In order to keep it simple, we will just discuss studies in which two variables are compared, but the idea is the same for more than two. Most commonly, you will be comparing your students' scores on different tests, such as an in-house proficiency test and the TOEFL. Some questions that might be answered by a correlational study are the folowing:

- To what degree does short-term memory relate to language proficiency?
- How well do the ratings on compositions by two different judges match?

- Do students who use more English during classroom activities score higher on the end-of-term test?

Note that in correlational research you are dealing with variables that exist more or less naturally, not ones that you manipulate. When you start manipulating variables, such as using pairwork with one class and TPR with another, you are moving into the realm of the experimental. Let us save that discussion for a bit later.

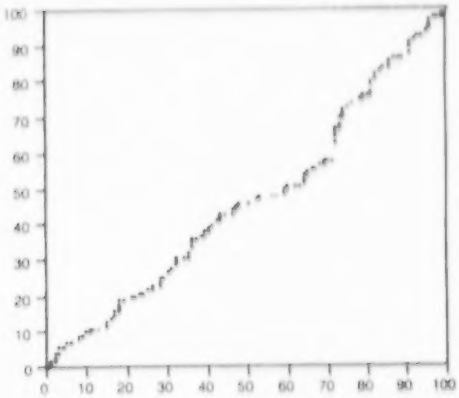
How do I set it up?

First, get a bunch of scores on one test. Then, get a bunch of scores on another test. Then, run the scores through a computer! Let us say that one day you notice that some of the students in your three-hour class seem more active after the mid-class break. You do some action research (or spying, as it is sometimes called) and observe the students during the break. You discover that half of them are going down to the local beer machine to have some refreshments before coming back up to class. So you ask yourself if there may not be some kind of relationship between drinking beer and participation. (See Guiora et al., 1972 and Guiora et al., 1980 for the effects of alcohol and other chemical substances on second language pronunciation ability.)

Next, you need to collect data. In our hypothetical example this would entail getting a "beer score" of how much the students are drinking and a "participation improvement score" showing how much more the students participate during the second half of the class.

How do I analyze the results?

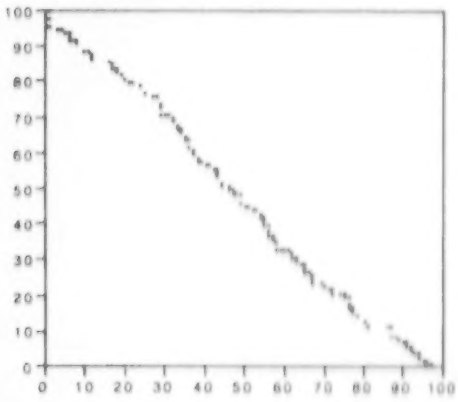
Once you get all the data together, you have two ways of looking at it: visually and statistically. If you just want to get a rough idea of what your relationship looks like, do it visually. Make a simple graph with the X-axis being the scores for one of the variables (e.g., the beer score), going from lowest to highest, left to right. The Y-axis would be the other variable (participation improvement score), once again going from lowest to highest, bottom to top. The results may look something like the graph on the left.



According to this graph, there is some kind of pattern which has developed. There is a fairly strong relationship between the two scores.

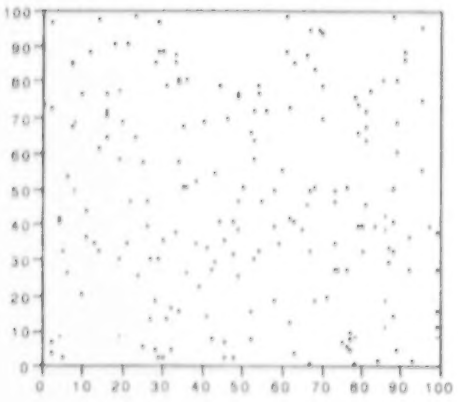
The left to right upward angle indicates that there is a *positive* degree of relationship here, and though the dots do not form a perfect line, it is pretty close. This means that, on the whole, the students who drank beer during the break actually did participate more during the second half of the lesson. But remember, that does not show that the beer caused the increased participation.

Let us assume, however, that the following graph resulted when the data were mapped out:



This graph also shows a fairly straight line pattern, but going from top to bottom. That shows a *negative* correlation, or an inverse relationship. In other words, as the drinking of beer goes up, the number of times a person has participated voluntarily has gone down in comparison to the participation for the non-beer drinkers in class. This means that one could predict that the more one drinks beer during the break, the less one would speak. Again, however, that does not prove that the beer was the cause of the decreased participation.

Now take a look at the graph in this figure:



That is a random plot reflected in that graph! There is no pattern at all. There is little or no relationship between the two variables. In other words, those who drank beer may have participated as much as, more, or less than those who did not. Therefore, it is very difficult to reach any conclusion about the relationship between beer drinking and classroom participation.

Calculating a correlation coefficient is slightly more complicated than using the eye-ball method, but it is much more en-

lightening as it can give you information on the strength of the relationship. If you are using a computer, just choose "correlation" from the menu and you are done. But because working with the statistical calculations is more complex, we do not have the space, time, wherewithal, or ability to go into further detail. Therefore we suggest that you move on to the appropriate chapters in this book, and take a look at the references listed below.

Where do I go from here?
As you can see, correlational research is quite a bit more involved than action research or using surveys. However, it is a very useful tool in analyzing relationships between different variables, and you should seriously consider making it a part of your research arsenal.

Experimental Research

What is it?
Experimental research is the oldest and most difficult type of research, and that is why we are not going to spend a lot of time confusing you about it here. Up to now, we have shown you that action research helps to solve everyday problems in everyday classrooms; surveys help you to discern patterns of responses; and correlational research helps you to discover relationships. Experimental research is for those times when "you have really got to know!"

You may wonder if three hours per week of listening will improve students' scores, as compared to the one hour per week you are teaching now. One may also ask if the amount of reading a student does has much effect on the student's ability to spell. Does the amount of time a student uses the Internet relate to the student's reading ability? In all of these cases, we are trying to see if there is some kind of cause-and-effect relationship between two variables which are well defined and easily measurable. To assess that relationship, the only thing to do, then, is to set up an experiment that isolates the two variables, get some data, and see how the data compare.

When do I do it?
Although the research purist may cringe, the ideal time for the average classroom teacher to set up and run an experimental study is . . . never. It is our belief that the true heavyweight research should be undertaken by trained professionals. Because of the difficulty and expense of doing a proper experimental study, most teachers would be much better served by action research or some sort of survey study. That is not to say that teachers are incapable of doing statistical research—we are just trying to be realists.

Having said that, we must now state that there is never a better time to get involved in hard experimental research than today! Teachers need to know where the researchers are coming from, and what better way to find out than by doing it

BEST COPY AVAILABLE

ourselves. One reason for the growing popularity of action research is that the teacher is a central player. His or her problems, concerns, or questions are being directly addressed. Experimental research can do the same, but only if the teacher gets involved. Indeed, some researchers have argued for greater involvement in research by everyday teachers (see Nunan, 1991; Johnson & Chen, 1992).

How do I set it up?

For now, the best way for teachers to get a general understanding of experimental research is to extend our earlier imaginary example. Recall that in the correlational study, the distinguished professor investigated the relationship between students' beer drinking during the break and participation after the break. Let us suppose that the professor found a strong positive correlation; in other words, the students who drank beer(s) during the break actually did participate more in the second half of the lecture. This led the professor to wonder if the beer really was responsible for the increased performance, or whether another unknown factor was at work.

First the professor randomly separated the class into two groups. One group, called the control group, would receive no special treatment. The second group became the experimental group, which would receive a measured amount of beer. The first half of the following lecture was held as normal, and a participation score was calculated for each "subject." Then, during the break, all of the students were herded into a different room where the "treatment" was given. The control group received nonalcoholic beer during the 15-minute break, but the experimental group received a predetermined amount of genuine beer. After the break, the class resumed as usual, and again the amount of participation for each student was recorded.

How do I analyze the results?

As we mentioned above, the goal in a typical experimental study is to determine whether a group that receives a treatment, such as a new way of teaching (or a controlled amount of alcohol), performs differently from a group that receives no special treatment. Unfortunately, an explanation of the statistics needed to analyze the data from an experimental study properly is beyond the scope of this chapter.

Where do I go from here?

If you are truly interested in experimental research, we urge you to pick up a couple of good books, such as the ones listed below or the ones mentioned elsewhere (see Griffee, this volume). Once you understand the concepts involved, get a grasp on the types of analyses necessary to do justice to your data, and come up with clear and precise research questions and hypotheses, you will be on your way. The time it takes to reach that point may be more time than most teachers are willing to invest, however.

One way to begin to get a grasp on experimental research is to have a look at Brown (1991) and Brown (1992). These articles try to help teachers understand

statistical research. They are not "how to do" articles, but more along the lines of "how to read and understand." Along with Brown's 1988 book, *Understanding Research in Second Language Learning*, they are essential reading for any teacher who has shied away from reading or carrying out experimental research.

Conclusion

We have attempted to introduce several types of research that may be of interest to teachers who, until now, have proclaimed a severe allergy to research. Although we have made light of research in this paper, we do not mean to imply that it is something that should be taken lightly. Every professional teacher needs to, at the very least, make an effort to understand research that is relevant to his or her teaching situation. We would also like to suggest that every teacher would be well served by trying out some type of research in his or her own classroom. Your goal does not have to be publication of a refereed article in a leading journal. Indeed your goal does not even have to be to write any article at all. But, all teachers should feel the need to improve themselves, and that is what classroom research is all about.

So, next time you have some trouble in your class, or next time a lesson bombs, or you are faced with students who do not seem to respond, or you hear about a great new lesson and wonder if it could work in your classes, stop wondering. Get out there and do something about it. Research may be the answer.

References

Ary, D., Jacobs, L. C., & Razavieh, A. (1990). *Introduction to research in education*, (4th ed.). Fort Worth, TX: Holt, Rinehart and Winston.

Brown, J. D. (1988). *Understanding research in second language learning: A teacher's guide to statistics and research design*. Cambridge: Cambridge University Press.

Brown, J. D. (1991). Statistics as a foreign language—Part 1: What to look for in reading statistical language studies. *TESOL Quarterly, 25*(4), 569–586.

Brown, J. D. (1992). Statistics as a foreign language—Part 2: More things to consider in reading statistical language studies. *TESOL Quarterly, 26*(4), 629–664.

Brown, J. D. (unpublished manuscript). Using surveys in language programs for needs analysis, program evaluation, and research.

Guiora, A. Z., Acton, W. R., Erard, R., & Strickland Jr., F. W. (1980). The effects of Benzodiazepine (Valium) on permeability of language ego boundaries. *Language Learning, 30*(2), 351–363.

Guiora, A. Z., Beit-Hallahmi, B., Brannon, R. C. L., Dull, C. Y., & Scovel, T. (1972). The effects of experimentally induced changes in ego states on pronunciation ability in a second language: An exploratory study. *Comprehensive Psychiatry, 13*(5), 421–428.

Hatch, E. M., & Farhady, H. (1982). *Research design and statistics for applied linguistics*. Cambridge, MA: Newbury House.

Johnson, D. M., & Chen, L. (1992). Researchers, teachers, and inquiry. In D. M. Johnson,

(Ed.), *Approaches to research in second language learning* (Chapter 9, pp. 212–227). New York: Longman.

Knowles, T. (1990). Action research: A way to make our ideas matter. *The Language Teacher, 14*(7), 7–10.

Nunan, D. (1991). Methods in second language classroom-oriented research: A critical review. *Studies in second language acquisition, 13*, 249–274.

Nunan, D. (1992). *Research methods in language learning*. Cambridge: Cambridge University Press.

Seliger, H. W., & Shohamy, E. (1989). *Second language research methods*. Oxford: Oxford University Press.

SECTION III

Research Methods

Chapter 7

Action Research: Something for Everyone

Gregory S. Hadley

Keiwa College

For many language teachers, the word "research" often evokes images of a controlled and scholarly project that is only understood by those in the mystic realms of linguists and second language acquisition research. Lightbrown (1985, p. 184) noted how in recent years researchers had "increasingly arranged their own research meetings, apart from the teaching conventions where they are always asked by someone in the audience to relate their findings to teaching practice." On the other hand, most teachers say they are either unable or unwilling to attempt their own research project, saying they lack the time, background, or experience to undertake such a project. Most teachers do not realize the benefits that come from doing small-scale action research projects. If they did, many would take pains not to miss such opportunities, and to invest themselves in the adventure of classroom-centered research.

The purpose of this article is to answer the following questions: What is action research? What obstacles must one overcome before doing an action research project? How can a new or inexperienced teacher begin an action research project? Finally, why should classroom practitioners attempt to do action research? After reading this article, language teachers will be prepared to start their own action research project.

What is Action Research?

During the past few years, attention has focused on one form of classroom-centered research (CCR) which is called Action Research (AR). LoCastro (1994, p. 5) defined AR as "one form of CCR which is seen as being small scale and situational . . . focused on a particular problem, to try to understand and perhaps solve some concrete problem in an individual teacher's classroom." She also suggested that AR should not to be done by outside researchers, but by actual classroom teachers. For a classroom practitioner who has never done an AR project, gaining the advice and assistance of an outside researcher or veteran teacher is quite beneficial. However, the classroom teacher should do the bulk of the work in order not to disrupt the natural flow of the class.

Kemmis and McTaggart (1988) also suggested that teachers should use AR with the goal of solving a specific problem in their classrooms. They defined AR as a process of planning, observation, and reflection on the part of the teacher with the goal of finding a solution to specific classroom problems. Other experts in the field (Cohen & Manion, 1985; Nunan, 1992) have written similar definitions.

I define AR as a process designed to improve teaching and facilitate learning by indentifying a specific classroom problem, targeting causes through systematic data collection (surveys, observation, interviews, etc.), and applying an effective solution to the problem as a result of the data being collected and interpreted.

Action research can be carried out by language teachers who do not have any special training in psychometric research methods or statistical analysis. It takes the pragmatic potential of research out of academic settings and puts it back into everyday classrooms. To use an analogy from medicine, action researchers are not specialists but rather general practitioners who diagnose and prescribe remedies for the everyday illnesses (low motivation, undefined learner goals, etc.) that language classes are prone to catch.

Obstacles to Action Research

Barriers to starting an AR project are: Vague research ideas, professional isolation, lack of available resources, and lack of time. These must be addressed before starting the research. The first step is to develop a clear research idea.

Developing a clear research idea may be difficult at first. Griffee (1994, p. 19) said that most research begins with a simple idea which might be noticing something or wondering why something is the case. Griffee added that another way to get a research idea is to pay attention to the areas of pain and frustration in your teaching. Problems are prime motivators of AR projects. As teachers, we sometimes need to "come up against a brick wall" in our classrooms before we are able to slow down and reflect upon what is not working well. Ideas for AR usually come naturally from this process of introspection.

If focusing on areas of difficulty or discomfort proves too daunting or metaphysical, teachers interested in AR should devote time to reading current teacher's journals that deal with problems, solutions, and issues that others have encountered in their classrooms. Sometimes such reading can reveal a problem that has previously gone undefined.

After identifying a problem in your classroom situation, write your research idea as a question that will help in finding a solution for the problem. It must be possible to find an answer to your question. For example, "Why do students come to class?" may be a vague and difficult question to answer, but the question "What classroom activities interest my students?" might be possible to answer through systematic observation, surveys, or other research tools.

Seek out other teachers while forming your research question. I believe that teachers improve their craft when in productive dialog with other professionals. If possible, join a language teachers' union or organization. In Japan, the most influential professional organization that language teachers can choose to join is the Japan Association for Language Teaching (JALT), but there are other organizations as well (JACET, Nippon Communication Gakkai, etc.). It is possible that a local chapter of a professional teachers' group meets in your city. If you do not have access to a local chapter, then seek out experienced teachers. Every city or region has at least one or two local "gurus" who are invaluable in helping new teachers get started. Find them. At all costs, break the isolation and start to network with other teachers. Share your ideas and pedagogic problems with them. New teachers are often surprised to find their problem is common to many in the profession. If this is the case in your situation, other teachers may ask to collaborate with you on your AR project.

After forming the research question and making contact with other colleagues, search for resources. It is vital to gain a better understanding of your research ideas. If you do not check the background of your subject properly, you may find out later that you have only repeated another person's work. Joining a teachers' organization such as JALT usually means you receive a regular magazine or journal full of interesting and helpful articles. Going to a regional or national language teachers' conference provides an opportunity to browse through publishers' displays of resource materials.

Some teachers may live near a prefectural or university library. However, as a word of warning, learning how to cooperate with a local library bureaucracy may require an investment of time before it bears much fruit. Having a computer with a modem is also a great tool for a literature search. Baskin (1994) listed contact names and addresses, current costs, and procedures involved in searching for journal articles and documents in Japanese universities and computer networks. Teachers wishing to do research in Japan would do well to read his article.

Time is probably the major concern of a teacher wishing to embark on an AR project. It is difficult to make the time for reading or networking. There are

always many immediate and pressing demands. However, the time spent on idea formation and reading for AR is an investment that produces lasting dividends. One's teaching and understanding of the processes involved in learning acquisition are enhanced by the insights gained from a literature search. The relationships we build through networking also enrich our lives. The increased rapport we gain with students who respond to our efforts can make a world of difference in the day-to-day grind of the school year. Not making the time for this type of endeavor is a lost opportunity.

Doing Your Own Action Research Project

Can a teacher with no previous experience undertake AR? The answer is an emphatic "yes." This section will outline an AR project I began with absolutely no knowledge of AR and which resulted in an encouraging turnaround in my classes.

The following points are paraphrased from Nunan's (1992, p. 19) seven-step cycle for AR:

1. Initiation	The teacher notices a problem in class.
2. Preliminary Investigation	The teacher spends time observing the class and taking notes of their behavior.
3. Hypothesis	After observation, the teacher forms a question or hypothesis as to the cause of the problem.
4. Intervention	The teacher tries several solutions to solve the problem.
5. Evaluation	After some weeks, the teacher consciously observes or measures the class again to see if there has been any improvement.
6. Dissemination	The teacher shares his findings with others.
7. Follow-up	The teacher looks for other methods to solve his original classroom problem.

Although I was not aware of Nunan's cycle when I started my project, I followed as similar process.

Initiation: Teacher Notices a Problem in Class

My technical college English class was not responding. From the first day of class, the students were slumping over their desks. They rarely looked up. When I greeted them, none returned the greeting, but instead looked furtively at each other before returning to their contemplation of their desks. Some slept through the class despite being awakened several times. To the simple question, "What's your name?", I several times received the response, "No." I was

assured by the school management that the students were merely shy and would eventually come out of their shells. However, the situation continued for weeks and then months. When no amount of class preparation seemed to work, I knew I had a real problem on my hands.

The first thing I decided to do was to begin networking. I joined my local chapter of JALT. It was there that I met experienced and concerned teachers who were willing to help me. Drawing on their advice, I found local libraries and requested catalogs from publishers of educational materials. I read texts that provided me with a better understanding of second language teaching methodology and theories. It was my hope that research could somehow help me find some solutions to my problem.

Preliminary Investigation

I read an article in *The Language Teacher* (Kobayashi, Redecop, & Porter, 1992) describing an AR project which examined Japanese college students' motivation to learn English in relation to university entrance exams. The findings were interesting for me because one of the groups that was surveyed was from a technical college. The authors found that while the interest in English-speaking cultures was high amongst the students surveyed, their intrinsic motivation—by which I mean motivation that is not attached to external rewards or incentives (see Deci, 1975, p. 23 for a more detailed discussion)—to learn English was much lower than their university counterparts. They speculated that perhaps lower scores on college entrance exams had something to do with lower intrinsic motivation but left the question open for further investigation.

Hypothesis: The Teacher Forms a Question After Observation

It occurred to me that my students, who had just recently graduated from high school, might have intrinsic motivation problems similar to those at the technical college students mentioned in the above article. I formulated the following question: Is the intrinsic motivation of Japanese technical college students to learn English lowered due to their lack of participation in or inability to pass university entrance examinations?

After forming my question, I began to read as much literature as I could find on the subject of motivation and the Japanese educational system. I quickly obtained books through publishers which helped me get a general idea about my topic. I also took advantage of the English research library at the Northern Illinois State University's satellite campus in Nakajo, Niigata. I took my time in reading and absorbed as much information as I could.

I then made a survey similar in focus to the one that Kobayashi et al. (1992) administered to their students (see Appendix 1). I omitted questions which did not pertain to technical college students, and added other questions designed to ascertain the students' intrinsic motivation to learn English. The survey was then translated into Japanese and distributed to 83 students.

Table 1
Comparison of Hadley 19913 and Kobayashi et al. 1992.
"NA" signifies questions not used in the survey.

Results of Survey (Percentage of Responses) Hadley 1993.	Results of Survey (Percentage of Responses) Kobayashi, Redecop, and Porter (1992).
1. Abroad to speak English Yes: 20% No: 80%	1. Abroad to speak English Yes: 4% No: 96%
2. Interested in foreign cultures Yes: 94% No: 6%	2. Interested in foreign cultures Yes: 83% No: 17%
3. Interested in explaining Japanese culture Yes: 53% No: 47%	3. Interested in explaining Japanese culture Yes: 52% No: 48%
4. Have a foreign friend Yes: 27% No: 73%	4. Have a foreign friend Yes: 19% No: 81%
5. Interested in English in Junior High Yes: 75% No: 25%	5. Interested in English in Junior High Yes: NA No: NA
6. Interested in speaking to a foreigner Yes: 100% No: 0%	6 Interested in speaking to a foreigner Yes: 73% No: 27%
7. Interested in English in High School Yes: 76% No: 24%	7. Interested in English in High School Yes: 60% No: 40%
8. Want to speak to foreigners outside school Yes: 27% No: 73%	8. Want to speak to foreigners outside school Yes: NA No: NA
9. English is important Yes: 75% No: 25%	9. English is important Yes: 87% No: 13%
10. English most enjoyable Kin: 2% Elem: 6% JH: 30% HS: 46% Pres: 16%	10. English most enjoyable Kin: NA Elem: NA JH: NA HS: NA Pres: NA
11. Took university entrance examination Yes: 18% No: 82%	11. Took university entrance examination Yes: NA No: NA
12. Examination facilitated speaking Yes: 19% No: 81%	12. Examination facilitated speaking Yes: 15% No: 85%
13. Watch English movies Often: 20% Some: 75% Never: 5%	13. Watch English movies Often: 8% Some: 50% Never: 42%
14. Class that motivates more Serious: 19% Relaxed: 81%	14. Class that motivates more Serious: 12% Relaxed: 88%
15. Regularly read other sources of English Yes: 17% No: 83%	15. Regularly read other sources of English Yes: NA No: NA

I tallied the results (see Table 1) and compared them to the percentages from the Kobayashi et al. (1992) study. I was surprised to find that my students were all (100%) interested in speaking to foreigners (question 6), and almost all (94%) were interested in foreign cultures (q. 2). Although the Kobayashi et al. study did not include a question about speaking to native speakers outside the classroom (q. 8), I added such a question because I felt that it would be a significant indicator of the students' intrinsic motivation. I noted that the same students (27%) who already had foreign friends (q. 4) were the only ones (27%) interested in speaking with a foreigner outside the school (q. 8). The low percentage indicated in question 8 was especially surprising since 20% of students in my sample had been abroad for the express purpose of speaking English, as opposed to the 4% in the Kobayashi et al. study. I was also shocked to find that only 18% of my students had even attempted to take a university entrance examination (q. 11).

The survey results were analysed using the *Survey Power 2.1* software package (Jepson, 1992). The responses to questions 5, 7, and 10 suggested that junior high school and high school were crucial times in the development of intrinsic motivation for studying English. The highest percentage (76%) of students said they were *interested* in English during their high school years (q. 7). Almost half (46%), felt that high school was the time when English study was *most enjoyable* (q. 10). This seemed to contradict everything that I had heard about the state of English language teaching in Japanese high schools.

Analysis of the results showed that 74% of the students who did not take an examination (q. 11) were also not interested in speaking to foreigners in English outside the classroom (q. 8); and, that the students who did not take a university examination were also not likely to read any English materials when not in school (q. 15). The implications of this correlation were not examined in detail, although, at the time my colleages and I suspected that non-participation in university entrance examinations may have influenced the students' intrinsic motivation to study English.

The survey suceeded in identifing certain times and conditions when our students were more interested in English, and in demonstrating that intrinsic motivation had been stronger before the students arrived at the technical college. Using the data from this survey, we began to think about some practical ways to try to revive that motivation.

Intervention: Offering Solutions

I discussed the results of the survey with the rest of the native English speaker (NS) staff at the school and noted the following: (a) a high percentage of students interested in speaking to foreigners within a "safe" environment such as the school or at school functions, (b) a strong student interest in English movies, and (c) a desire—amongst the students—for a more relaxed classroom setting. Taking advantage of my role as head teacher at the school, other NS

teachers and I attempted strategies we hoped would raise the students' intrinsic motivation to learn English.

More native English speakers were invited to our seasonal school parties. In the past, the parties had been a closed affair which looked sadly like a middle school dance party, but students began to interact with the native speakers and use the opportunity to speak English in a natural setting. Together with other NS teachers at the school, we introduced video to the classroom, using focused listening and information gap activities. We knew how important pop music was to our students, so we incorporated a number of activities using music (see Griffee, 1992). To foster a more relaxed classroom setting, we changed our old texts which were based on an audiolingual syllabus and switched to a more communicative-based syllabus. We included many task-based classroom activities to get students out of their chairs and speaking to other group members. Borrowing ideas from Moskowitz (1978) as well as Davis and Rinvolucri (1990), we used more "caring and sharing" activities so that students could express themselves on issues important to them, such as dating, jobs, and entertainment. We encouraged them to work together in their language learning.

On occasion, we took our students out of the classroom and taught in parks, markets, and coffee shops. Although not all the students warmed up to it, we instituted a drama day at the school where students could exercise their skills in theater and role play.

Evaluation

After three months, all teachers observed a higher level of classroom participation. Whereas students had often been absent, unresponsive, or asleep, we found that the majority were laughing, speaking, and even joking with the teachers in class. Several of the students made new friends or penpals with native English speakers outside the school, and the students began to have more conversations with teachers outside of class. After graduation, a number of students went on extended personal trips to Australia, England, and the United States for the purposes of meeting penpals, learning more about English-speaking cultures, and improving their language skills. We saw a dramatic turnaround in our students' motivation thanks, in part, to our efforst to apply what we learned through AR.

Dissemination: Sharing the Findings with Others

Teachers can share the findings of their AR projects in local teachers' meetings as a presentation, in informal meetings with other colleagues, or by publishing their results for a larger body of readers. While many teachers seem to find articles written by linguists to be dry and unrelated to their classrooms, the same teachers are quick to listen to the findings of a colleague who shares the same day-to-day difficulties. Action research reports get read and appear to have

greater immediate impact on the practices of other classroom teachers than the findings of second language researchers.

Follow-up: New Solutions

Out of curiosity, I networked with other teachers at technical colleges. I wanted to see if the survey results from my school would be similar if we took a larger sampling from other schools in Niigata prefecture and the Tokyo metropolitan area. We administered the same survey to 562 students. The results were indeed similar, yet not as clearly delineated as in the original survey (see Appendix 2). This follow-up of my research on a larger scale lent credence to my earlier findings, and helped me to show my colleagues that the dynamics present in my school might also be present in theirs, hopefully spurring them on to beginning AR projects of their own.

My contract reached its two year limit, and I moved on to teach at the college level. If I had had more time at the technical college, I would have looked for solutions in addition to the ones we tried. This is another aspect of AR: The process is constantly in motion. One cannot afford to rest on one's past accomplishments, but must move forward to keep on the cutting edge of the class's task of learning.

Why Should Teachers Do Action Research?

The advantages of having AR in one's "teaching toolbox" are clear. Apart from improving one's craft, teachers will gain a greater insight into what is going on in the minds of their students. Clearer communication will be fostered between teacher and student. Few teachers work well in a professional vacuum, so opportunities for networking with other dynamic and conscientious educators is a definite plus.

Instead of being a reactive teacher who fumbles in the darkness, groping for something that will work in class for that day, one can become a pro-active teacher through the thoughtful use of AR. Based upon the data gained from research, a teacher can seize the moment and move forward with purpose and clarity.

Conclusion

When I arrived in Japan over three years ago as an inexperienced teacher, I encountered various problems. Not only did I feel isolated and overworked, I also had to deal with culture shock. At the same time, I was struggling to find ways to reach a particularly gloomy, poorly-motivated group of technical college students in a dilapidated building. The administration had no training program for new teachers. Our teaching materials consisted of tattered 15-year-old TESL textbooks obtained from what appeared to be an American college bar-

BEST COPY AVAILABLE

gain basement sale. I was overwhelmed and didn't know where to go for help. It is in this type of environment that AR flourishes.

References

Baskin, R. (1994). Research in Japan. *The Language Teacher 18* (11), 14–15, 110.

Cohen, L., & Manion, L. (1985) *Research methods in education*. London: Croom Helm.

Davis, P., & Rinvolucri, R. (1990). *The confidence book*. Essex, UK: Longman.

Deci, E. L. (1975). *Intrinsic motivation*. New York: Plenum Press.

Griffee, D. (1994). The art of language research. *The Language Teacher 18* (2), 19–21.

Griffee, D. (1992). *Songs in action*. London: Prentice Hall.

Jepson, T. (1993) *Survey Power 2.1 for DOS*. Wisconsin Rapids, WI: Wisco Computing Company.

Kemmis, S., & McTaggart, R. (Eds.). (1988). *The action research planner* (3rd ed.). Geelong, Australia: Deakin University Press.

Kobayashi, S., Redecop, B., & Porter, R. (1992). Motivation of college English students. *The Language Teacher 16* (1), 7–9, 15.

Lightbrown, P. M. (1985). Great Expectations: Second Language Acquisition Research and Classroom Teaching. *Applied Linguistics 6* 184.

LoCastro, V. (1994). Teachers Helping Themselves: Classroom Research and Action Research. *The Language Teacher 18* (2), 4–7.

Moskowitz, G. (1978). *Caring and sharing in the foreign classroom*. Rowley, MA: Newbury House.

Nunan, D. (1992). *Research methods in language learning*. Cambridge: CUP.

Appendix 1

Survey

1. Have you been abroad to speak English?
   [ ] A. Yes
   [ ] B. No

2. Are you interested in foreign cultures?
   [ ] A. Yes
   [ ] B. No

3. Are you interested in explaining Japanese culture to foreigners in English?
   [ ] A. Yes
   [ ] B. No

4. Do you have any foreign friends?
   [ ] A. Yes
   [ ] B. No

5. Were you interested in studying English in junior high school?
   [ ] A. Yes
   [ ] B. No

6. Are you interested in speaking to a foreigner?
   [ ] A. Yes
   [ ] B. No

7. Were you interested in studying English in high school?
   [ ] A. Yes
   [ ] B. No

8. Do you ever try to speak English with foreigners outside the school?
   [ ] A. Yes
   [ ] B. No

9. Do you think that English is important for your future job?
   [ ] A. Yes
   [ ] B. No

10. When did you enjoy English the most (circle one)?
   [ ] A. Kindergarten
   [ ] B. Elementary School
   [ ] C. Junior High School
   [ ] D. High School
   [ ] E. Present School

11. Did you take a University Entrance Examination?
   [ ] A. Yes
   [ ] B. No

12. Did the university entrance examinations help you to speak English?
   [ ] A. Yes
   [ ] B. No

13. How often do you watch English movies or videos?
   [ ] A. Often
   [ ] B. Sometimes
   [ ] C. Never

14. Which type of class motivates you more, a serious or relaxed class?
   [ ] A. Serious
   [ ] B. Relaxed

15. Do you regularly read books, magazines or newspapers in English (not including your textbook)?
   [ ] A. Yes
   [ ] B. No

Appendix 2

Motivation of English Technical College Students: Results of Survey (Percentage of Responses) Hadley, Fountaine, and Megill 1993.

1.	Abroad to speak English	Yes: 44%	No: 56%
2.	Interested in foreign cultures	Yes: 90%	No: 10%
3.	Interested in explaining Japanese culture	Yes: 57%	No: 43%
4.	Have a foreign friend	Yes: 38%	No: 62%
5.	Interested in English in Junior High	Yes: 68%	No: 32%
6.	Interested in speaking to a foreigner	Yes: 93%	No: 7%
7.	Interested in English in High School	Yes: 65%	No: 35%
8.	Want to speak with foreigner	Yes: 37%	No: 63%
9.	English is important	Yes: 79%	No: 21%
10.	English most enjoyable	Kindergarten: 7%	
		Elementary School: 8%	
		Junior High: 21%	
		High School: 45%	
		Present School: 23%	
11.	Took university entrance examination	Yes: 31%	No: 69%
12.	Examination facilitated speaking	Yes: 33%	No: 67%
13.	Watch English movies	Often: 29%	
		Sometimes: 64%	
		Never: 7%	
14.	Class that motivates more	Serious: 23%	Relaxed: 77%
15.	Regularly read other sources of English	Yes: 20%	No: 80%

Chapter 8

Students as Ethnographers

Linda Donan

In a busy Osaka fast food restaurant at lunch time, three Japanese junior college students surreptitiously position themselves to watch a foreign couple eating. One student takes note of how often the woman uses her napkin compared to the man; another student counts the times the two customers reach across the narrow table to touch each other during their conversation; and the third student uses two stopwatches, a pencil, and paper to record how much time the couple spend in verbal exchange and in silence. As soon as the couple finishes their lunch, the students select an entering Japanese couple who are apparently about the same age as the foreign couple, and set up to resume their surreptitious observation.

Meanwhile, in Mie prefecture, an American researcher has arrived at a small fishing village. He is being formally introduced to the mayor by a go-between who is also a researcher at a respected Japanese University. Later, the mayor will urge various village members to allow the researcher to live for extended periods in their homes. After a year or more, the researcher will have sufficiently ensconced himself into the village life, to report on the modern "house-husband" role attributed to the partners of traditional *ama* (female pearl divers).

Both the professional researcher, dedicating a year or more of his life, and the college students giving up their lunch hour, are practising ethnographic research.

Ethnography

The word ethnography is formed from the two Greek words "ethnos" (tribe or nation) and "graphos" (something written). Both the researcher and the college students are studying members of a tribe or nation. Both the professional researcher with his suitcase full of notebooks, cassette tape recorders, word pro-

cessor, video, and time-lapse cameras; and the students with their simple timers, notebooks, and pencils are doing their ethnographic study by writing something down or recording what they observe.

The ethnographer, Clifford Geertz (1973) noted that it is the act of inscription which transforms momentary events into accounts that can be reconsulted.

If anyone should ever wish to know who washes the dishes in the home of the *ama*, or if anyone wants a glimpse at the eating habits of foreigners in fast food restaurants in Japan, they can consult a written report.

Miles and Huberman (1984) pointed to the use of ethnography in providing rich descriptions of local communities. The pearl diving village described above is a local context chosen for its unusual characteristics. For the purpose of cross-cultural or comparative research, however, a field such as the fast food restaurant selected by the college students is richer because people of many cultures can be found performing the same ritual—eating their hamburgers.

Frey (1988) contrasted the subjectivity of ethnography with the objectivity of fact finding. Although there are tools of measurement used in both of the examples described above, the most important tools are the observers themselves. The ethnographic investigator looks at a subject that they are ideally equipped to understand, namely, human beings. Observing actions, listening to discourses, and in longer studies discussing the subjects' perceptions of themselves are the primary ways that data is gathered for ethnographic research.

Ethnographers also differ from experimental researchers in their relationship to the actions they observe. While the latter generally attempt to maintain a personal distance from the observation, ethnographers often interact closely with their subjects or with the subjects' environment. Lindlof (1988) extended this notion of involvement to include ethnographic descriptions that require the richness of the observers' experience in the field to give fullness to the situation being described.

Why Choose Ethnography?

Experiments, or other forms of quantitative research, are useful when working with numerically measurable data under controlled conditions. They provide excellent techniques for getting very specific information. They are not, however, always appropriate for working with human subjects. People often act in complex and unexpected ways that are beyond the parameters of an experimental hypothesis. Experimental exactitude, which might require controlled conditions, tends to make people behave self-consciously or unnaturally.

Ethnographers and other qualitative researchers try to get a wider view. They trade specificity and measurability for the experiential depth of watching what people do in their normal activities in their normal life spaces.

Qualitative and quantitative research are often viewed as polar opposites appropriate for different tasks or different departmental paradigms (i.e., quanti-

tative research for the hard sciences and qualitative research for the social sciences). It is possible that some types of exploration may best be carried out through a series of research strategies beginning with ethnographic research and progressing through to systematic quantitative studies. The choice to be an ethnographer rather than an experimental researcher can also involve issues of personal conviction based on personality traits, attitudes, and values.

What are the Steps of Ethnography?

Careful planning and procedures can help to insure a positive outcome to ethnographic studies. Some steps you may want to consider are:

1. Preplanning
   (a) Choosing a study place (field) or subject (tribe).
   (b) Determining limitations and possible problems.
   (c) Casing the scene for feasibility and suitability.
   (d) Making your formal research proposal.
   (e) Gathering tools and arranging logistics.
2. Getting in and getting help
   (a) Protecting human subjects.
   (b) Convincing gatekeepers to allow access.
   (c) Making use of a "friendly native."
3. The actual observation
   (a) Choosing a role.
   (b) Taking field notes and processing them.
   (c) Interviewing the observed.
4. Follow-up
   (a) Analyzing your notes.
   (b) Developing concepts.
   (c) Interpreting and evaluating.
   (d) Writing up and publishing results.

The above list is not to be considered as an inflexible procedure, but rather as a set of useful choices. Each ethnographer must follow the steps that prove successful—through trial and error—with their own subjects, in their own fields.

1. Pre-planning considerations

(a) *Choosing a study place (field) or subject (tribe).* Determining which cultures and which subjects to study is often based on personal interest and opportunity. The Mie researcher described earlier in this paper, came from a society where shared household duties is a modern norm to study this marginal phenomena in a society where household roles are usually divided between the sexes. For the college students a sentence in an English textbook about the

fastidious eating habits of Japanese people prompted their study. Their teacher had read earlier research results comparing touching customs in various cultures and had suggested that research question to them.

Having decided on the subjects, one must then choose the field or place of observation. For college students, unable to travel abroad, the local fast food restaurant provided an accessible site for ethnographic research. However, if foreign restaurant-goers never entered that restaurant then it obviously could not be chosen as a field of research. If lighting was low, tables were curtained off, or if the students could not get a table they would have had observation difficulties.

(b) *Determining limitations and possible problems*. The tribe and field one chooses to study will produce unique problems that must be considered.

The researcher's own limitations must also be considered. The Mie researcher had considered his own gender, his height, his clothing, and obvious foreignness and taken into account how those factors may effect his interaction with the subjects. He had prepared himself by studying the Japanese language, but had also taken into account his understanding of Japanese culture and potential ethnocentrism or prejudices that might color his observations. Likewise, his own positive attitudes towards males doing housework needed to be seen as a possible problem that might affect his findings.

The college students' inexperience at research was compensated for by the simple numeric techniques they chose. Nevertheless, one of the students revealed the effects of ethnocentrism on observation when testing her hypothesis that the Japanese subjects would use their napkins more often than the foreign subjects. She had equated napkin use with a personal value of cleanliness. Her results, however, contradicted her hypothesis, a fact which she found so personally and emotionally upsetting that she re-did the observation several times. Her co-researchers had to ask her to be silent in order to prevent her incredulous and angry remarks being overheard by the subjects under observation.

(c) *Casing the scene for feasibility and suitability*. The next step, casing the scene, does not necessarily start at the scene. If at all possible, ethnographers should consult the literature for previous or on-going studies of their subjects, and consider the recommendations of experienced people before they visit the field. A researcher studying African tribes was once given the unexpected advice, by an experienced ethnographer, to carry a stick to ward-off dog attacks (Sanjek, 1990).

Colleagues of the Mie researcher had told him of the *ama* and their domestic arrangements, and he had read anecdotal reports as well as Japanese research studies. He also read about mainstream Japanese attitudes towards housework, and discussed the issues involved with his Japanese acquaintances. The college students had discovered that only the one of them that was sitting nearest the table under observation could hear well enough over the noise to be certain whether the subjects were speaking or not. They realised, however, that they could sometimes watch the subjects' lips either directly or reflected in a nearby wall mirror.

(d) *Making your formal research proposal.* For many ethnographers, the next task involves setting up a formal research structure. The Mie researcher had to submit a research proposal to get time off from his university teaching. He also had to convince the funding board of his rationale for the study. The college students had received their teacher's verbal approval of their project, and they had arranged a meeting time and decided who would do which part of the observation.

(e) *Gathering tools and arranging logistics.* Equipment needs can also be determined and prepared in advance. Traditionally, an ethnographer's tools have included two eyes, notebook, and a pencil. Recently, recording and counting devices have aided observation, but one should be careful to take into the field only those tools that can be handled easily while focusing attention on the subjects.

If possible, the observer should conduct a pilot study or informal visit to a proposed site to determine such things as the best time to conduct the observation, and the best vantage point. Earlier studies in Mie by Japanese anthropologists focusing on the *ama* had prepared the villagers for the inconveniences of tape recorders and strangers in their homes. For the student observers, the lunch hour rush proved an ideal time to observe subjects from a well centered table that afforded a view of all other tables.

2. Getting in and getting help

(a) *Protecting human subjects.* Care must be taken that the subjects of ethnographic research be allowed to go about their natural lives without feeling imposed upon by the observation process.

In America, where experimental studies in bygone decades allowed researchers to shock their subjects with electrodes, strict measures are now in place to protect human research subjects. The American researcher, therefore, had to fill out legal documents to be considered by a board in charge of the safety and happiness of human research subjects. Throughout his stay, he had to ask again and again if people minded him sitting in on their daily activities, and request permission for all his intrusions including the use of tape recorders and video cameras. Any slight hesitancy or protest on the part of his subjects required him to retract and wait until the people were more comfortable with his presence, or to withdraw to another home.

No such standards were set for the students in Japan. Their teacher did suggestion that the content of private conversations not be recorded, and in another case from the same class the teacher suggested that the interviewed foreigner not be approached for a "date."

(b) *Convincing gatekeepers to allow access.* Getting into the research field often requires admission by a "gatekeeper." The Mie researcher was introduced by a prominent professor affiliated with Tokyo University. The Mayor also worked on his behalf securing the confidence of local people. These go-betweens helped to smooth the way for the study. The students could have asked for permission

from the restaurant's employees and manager but chose to remain incognito—even to their gatekeepers—and to rely on the fact that one of them was a former employee to get them out of any trouble if it arose.

(c) *Making use of a "friendly native."* Japanese people are well aware of the uses of go-betweens as gatekeepers and can accept the need for a "friendly native" to help with translating not only the language of the subjects but also the culture. Especially in observing a culture or sub-culture other than one's own, the help of a friendly native can save the researcher from errors such as that of the ethnographer, Margaret Meade (Howard, 1984). Meade did not find out until years later that the subjects of her south sea island research had been "pulling her leg" with their sexual stories because she seemed to enjoy such tales. She did not have a friendly native who was frank enough with her to explain that her research was being skewed by the imaginative stories that they thought would please her.

3. The Observation

(a) *Choosing a role.* Ethnographic researchers must choose a role to play in the field. Some possible roles are: complete participant, participant-as-observer, observer-as-participant, and complete observer. The college students pretending to be customers and hiding their pencils behind their lunch were aiming at the complete participant role. This role can be particularly effective in that subjects are usually unaware of being observed. It is also effective in that the researcher shares the experience of the subjects, and knows, for instance, if the greasiness of the fries makes one want to use a napkin or if the noise level of the restaurant discourages conversation.

The work of Trujillo and Dionisopoulos (1987) observing policemen's use of "labeling" and "talking tough" by riding along with them in the front seats of patrol cars is an example of participant-as-observer. They were known to be conducting research and were not wearing police uniforms. They were not expected to take an active part in arrests, and could completely withdraw from dangerous situations. They negotiated a space to cohabit, and the police never expected them—nor would allow them—to "face the heat."

Participant-as-observer was also the role chosen by the Mie researcher. He could not pass as a Japanese nor become a second house-husband in the situation. Although he hoped to be granted the comfort of being included in the family over time, he remained aware of the limitations of his identity as a guest, and maintained an awareness of his obtrusiveness.

An example of the observer-as-participant role is the interview. In this situation, the researcher is on the scene for a brief period and clearly playing the role of observer. Their clipboard is their badge that allows them to stop complete strangers on the street and take up their time with questions. Researchers who do not make their role clear, for instance during random street interviews, might receive hostile reactions, be regarded as crazy, or even be reported to police.

To be a complete observer, in other words to see and not take part or be seen, is possible these days with the use of cameras and tape recorders. Many social scientists dislike this method of gathering data as it can stretch the rules of privacy and it does not allow the researcher enough contact with his or her subject. Clearly, the categories are not simple or definite but rather there is a spectrum of possibilities. The Mie researcher may have pitched in and washed dishes with his house-husband hosts, and the students I am sure took some sips of their shakes and bites of their hamburgers.

(b) *Taking field notes and processing them.* Taking field notes is the heart of ethnography. Questionnaires and taped interviews can be quite comfortable tools for researchers interacting in a culture and language that is not their own. The class of the Japanese college students in the above-mentioned class in Osaka had learned that neither of these tools were as effective as observation, for their particular questions. A questionnaire with a question such as, "How often do you use a napkin when you eat?" elicited responses ranging from nervous and perhaps less than honest "very often" replies to the more honest but less helpful "I don't know, I don't count."

Field notes are ideally massive and detailed notes taken while the observation is in process or at the first available free moment unobserved in the case of complete participants or participant-observers. A detailed observation in the Mie village might be something like:

> 5:08, Wednesday morning, week 23, the Tanaka husband enters the kitchen from the bedroom barefoot in his wrinkled pyjama bottoms. He is the first person to leave the bedroom this morning. He crosses the floor in three shuffling steps, turns on the rice cooker sitting on the table with a press on a single lever and yawns widely. Husband notices I am awake and smiles widely. Says "It's early, isn't it." I agree with a nod. Husband scratches self in the right ribs. Husband nods four deep nods, bows in my direction, crosses the kitchen to the bedroom door and apparently returns to bed.

Such detailed notes might be accompanied by a floor plan of the Tanaka kitchen. They might be full of abbreviations such as "TH" for Tanaka Husband, "4 n, 1 db" for four nods and one deep bow. Also, good field notes contain parenthetical-notes such as (wide-smile appears to be one of discomfort rather than friendly recognition), (all family members practice shuffling steps when anyone in the house is presumed to be asleep), (right ribs seem to be the preferred spot of self-scratching of TH whereas left ribs seem to be the preferred spot of nervous self-comforting of Tanaka son, daughter, and visiting cousin).

The teacher of the Osaka college students had difficulty teaching her students to differentiate between true observations and conjectures, a problem which often resulted in incomplete notes. Students had to be shown again and again to place parentheses around their own opinions and to keep them separate and treat them

differently from the observed data. "He wiped his hands on his pants" is observation, whereas "He is unclean" is parenthetical opinion. Also, parenthetical statements had to be expanded back into observations. The observer wrote, "She was happy when he touched her hand," to which the teacher asked, "How exactly did you know she was happy? Did she smile? Did she lean forward? Did her eyes narrow or widen? Did her breathing or facial coloring change? Did she say, 'I am happy?'" Ethnographic observation, like all skills, improves with practice.

Tape recorders and video cameras are thought by some to have eased the necessity for detailed field notes. But there are fields of study which lack both available electricity or a supply of batteries for outdoor observers, and there are fields where movement, action, concealment, and privacy make these devices too unwieldy. Even when a video-camera has been in place, however, the observer's field notes are necessary. The human neck swivels to catch the smallest peripheral action, the human eye moves to change focus much more reactively with the human event than a camera on a tripod. The discomfort that subjects feel when under the camera's eye and how that effects the naturalness of their behaviors is also an important consideration.

One teacher studying classroom action in Nagoya gets around the problem of subject awkwardness by bringing two video cameras to class from the beginning of the semester. Students are asked to ham-up their self introductions, which are videotaped. Toward the end of the collegiate year, apparently, the students are so desensitized to the camera's presence that the teacher can zoom-in on the most private of pairwork conversations with impunity.

During fieldwork, the Mie researcher often perused his field notes to see if some patterns of behavior or communication where forming or if some interesting and unusual paths of inquiry have been overlooked. In the case of the shorter studies, a researchers can use their field notes to find points worthy of a further study or areas they have neglected to complete in the current study.

4 Follow-Up

Evans-Pritchard (1951), who did anthropological studies in foreign countries such as Sudan under the most rustic and difficult of conditions in the 1950s, felt that the most difficult aspect of ethnographic study was writing-up the notes. Certainly injuries, illnesses, and other discomforts are part of the quality of being in the field; but it is in the writing-up of field notes that real mental difficulties arise. Not only is it a gargantuan task, but unlike the social task of being out and observing your subjects, it is a lonely one usually spent at a desk with your field note journals piled around your elbows. One cannot merely type it up and produce a treatise on what happened and then what happened next. To make field notes readable the researcher must dig through the mountains of notes or tape recordings, mine out the valuable insights, and present them in a natural setting. The "mining tools" are the researcher's own deep understanding of the observed tribe.

Are There Special Considerations for Ethnography in Japan?

If you are not a native Japanese, observing Japanese subjects has a possible problem with cultural misunderstanding and thus misinterpretation in the research. Even if you are fluent in the Japanese language, there may be some cases where interviews or overheard conversations are not clear to you. So much of the fashion and customs of my own country, the USA, can be seen here in the large and modern cities of Japan that it is easy to forget that they are a facade over a nation that is unique in values, attitudes, and heritage. I believe it may actually be easier for an ethnographer working in a mountain village in Papua New Guinea to keep in mind the basic rule of "Watch and listen carefully, for you are a stranger here and anything and everything is important!" In Japan, where things seem to be easy to understand and one does not need to carry a big stick to fend off dogs, the observer may become less cautious than an ethnographer must be. Too many important points may be missed.

Here, too, one must make even greater efforts to ensure that a friendly native keeps an eye on how your observations are interpreted. Many Japanese people, however, who are friendly enough to help with your ethnography will not be culturally disposed to frankly criticize your mistakes or misunderstandings.

There are, also, specific ethical considerations in this field. Many foreigners in Japan are here as English teachers and using your students as reporters and research-aids is an understandable choice. But, one must question whether taking class-time to teach ethnography techniques and to process interviews and observation data is an appropriate use of English study time. The students in the Osaka example above had signed up, not for a regular English conversation course, but for a special course titled "Intercultural Communications and Research."

Japan does not, apparently, have laws concerning privacy and protection of humans in experimental studies. Without the rigour of a review board, you must be more than conscientious in the way you set up, practice, and later publish your research. The ethnographer works with people, and must never forget that people can be harmed.

Conclusion

There are just about as many ethnographic styles of research as there are ethnographers, and one must choose a method and refine it in the field to fit one's own nature and circumstances. All expatriates in Japan and their students may really find it fun to look around them and see all that is to be seen in this tribe. The follow-up publishing of one's ethnography in computer postings, professional journals (see Appendix), or in film documentaries must be done with ethical considerations but it can be fun as well.

References

Evans-Pritchard, E. E. (1951). *Social anthropology*. New York: Free Press.

Foley, D. E. (1990). *Learning the capitalist culture: Deep in the heart of Tejas*. Philadelphia: University of Pennsylvania Press.

Frey, L. R. (1988, November). *Meeting the challenges posed during the '70s: A critical review of small group communication research during the '80s*. Paper presented at the meeting of the Speech Communication Association, New Orleans. In Frey et. al. (1991). *Investigating communication: An introduction to research methods*. Englewood Cliffs, NJ: Prentice Hall.

Geertz, C. (1973). *The interpretation of cultures*. New York: Basic Books

Gephart, R. P., Jr. (1988). *Ethnostatistics: Qualitative foundations for quantitative research*. Newbury Park, CA: Sage.

Howard, J. (1984). *Margaret Meade: A life*. New York: Fawcett Crest.

Lindlof, T. R. (1988). Media audiences as interpretive communities. In J. A. Anderson (Ed.). *Communication yearbook 11* (pp. 81–107). Newbury Park, CA: Sage.

Miles, M. B., & Huberman, A. M. (1984). *Qualitative data analysis: A sourcebook of new methods*. Newbury Park, CA: Sage.

Sanjek, R. (Ed.). *Fieldnotes: The makings of anthropology*. Ithaca, NY: Cornell University Press

Trujillo, N., & Dionisopoulos, G. (1987). Cop talk, police stories, and the social construction of organizational drama. *Central States Speech Journal, 38*, 196–209.

Appendix: *Some Journals that Publish Ethnographic Studies*

Anthropology/Sociology	Communication	Education/Language
American Anthropologist	*Communication Theory*	*Anthropology and Education Quarterly*
American Ethnologist	*Critical Studies in Mass Comm*	*Discourse Processes*
Anthropological Quarterly	*Cultural Studies*	*International Journal of Qualitative Studies in Ed.*
Cultural Anthropology	*European Journal of Comm.*	*International Journal of Sociology of Language*
Current Anthropology	*Journal of Comm. Inquiry*	*Language in Society*
Ethnology	*Media, Culture & Society*	*Public Culture*
Family Process	*Research on Language & Social Interaction*	
Human Studies	*Semiotica*	
Journal of Contemporary Ethnography	*Signs*	
Journal for the Theory of Social Behavior	*Studies in Visual Comm.*	
Journal of Applied Anthr.	*Symbolic Interaction*	
Man	*Text & Performance Quarterly*	
Qualitative Sociology		
Sociological Quarterly		

Chapter 9

Designing Surveys for Language Programs

James Dean Brown

University of Hawaii at Manoa

A language survey is any procedure which gathers information about participants' characteristics and views on the nature of language or language learning through the use of oral interviews or written questionnaires. In this paper, I will begin by examining the different types of surveys, including both interviews and questionnaires. I will then discuss how language surveys can be used for language curriculum development and research. I will continue by exploring the specific functions that surveys can perform. The next two sections will provide a set of guidelines for writing effective survey questions along with discussion of the issues involved in achieving a good return rate. I will then present a brief section on analyzing survey results and will end the paper with suggestions for further reading.

Types of Surveys

As shown in Table 1, surveys can take two basic forms: interviews (including individual, group, and telephone interviews) and questionnaires (including self-administered or group administered questionnaires). Each of these categories and subcategories merits a bit more discussion.

Table 1
Types of Survey Procedures

Type of Survey	Procedures
Interviews	Individual
	Group
	Telephone
Questionnaires	Self-administered
	Group-administered

Interviews

Interviews are procedures for gathering information orally in specific, planned categories (if the interview is well structured in advance) as well as information that was not anticipated at the outset. Interviews can be conducted in three ways: individually, in groups, or by telephone.

Individual interviews are interviews conducted with participants on a one-to-one bases. Such interviews gather information privately, so a certain level of confidentiality can be maintained. As a result, individual interviews have the advantage (over group interviews) of collecting something close to the true views of the participants. However, such interviews also have the disadvantage of being relatively time-consuming. As a result, individual interviews may be most effective for finding out which questions, views, or issues are worthy of later follow-up study using more structured group interviews or questionnaires, both of which can be administered much more efficiently than individual interviews.

Group interviews, which are interviews conducted in a group or meeting format, can be used to circumvent the time-consuming nature of individual interviews. However, group interviews also have the disadvantage of lacking confidentiality. For personal, political, or emotional reasons, the opinions expressed when participants are in a group interview may be quite different from the views expressed when they are interviewed individually and confidentially.

Telephone interviews sometimes provide a useful compromise in that they are confidential and somewhat less time-consuming (saving the time necessary to travel to the participants). However, there are also several disadvantages that must be considered in using this type of interview: (a) in a telephone interview, some information may be lost because the interviewer cannot see the facial expressions, gestures, and surroundings of the participants; (b) a telephone survey also precludes the possibility of showing participants anything during the interview; and (c) at least in the United States, many people have become wary of telephone sales promotions and surveys and may not cooperate.

Questionnaires

The various types of interviews can clearly be used to explore the questions, topics, views, opinions, etc. of participants in a particular language teaching institution. However, if these points are worth pursuing on a larger scale, questionnaires may prove much more efficient. A *questionnaire* is any written instrument that presents participants with a series of questions or statements to which they should react either by selecting from existing possibilities or writing out their answers. Questionnaires are particularly efficient for gathering information on a large scale. For example, an interview with a few of the participants in a language program might be helpful in designing a questionnaire, but the questionnaire itself would be much more practical to administer to all of the participants. Most questionnaires are the self-administered type. A *self-administered questionnaire* is usually mailed out and filled out by the participants in their own homes or offices (i.e., they are self-administered) and then returned by mail. Self-administered questionnaires have three potential problems: (a) they often have a very low return rate, (b) they must be completely self-explanatory because further clarification is not possible, and (c) they are administered under conditions that are unknown to the survey designers. A *group-administered questionnaire* is any questionnaire which is administered to participants while they are in class or at a meeting. For example, if I wanted to survey a particular group of students, I could go to their classes and ask their teachers to let me use the 15 to 20 minutes necessary for the students to fill out my questionnaire. By using such a strategy, I would solve all three of the problems listed in the previous paragraph for the self-administered questionnaires: (a) the students will be a captive audience and will feel obliged to fill out the questionnaire (thus giving me a high return rate), (b) I can be present to explain any ambiguities as they arise, and (c) I will know exactly what conditions existed when the questionnaires were filled out.

Uses for Language Surveys

Surveys are most commonly used in language education for curriculum development and for research. Language programs provide a service to students who want to learn the language in question. If the program is interested in what these *clients* think, then some form of curriculum analysis and research will probably be desirable because the best way to find out what people are thinking about any aspect of a language program is to ask them.

Surveys for Curriculum Development

The two elements of curriculum development wherein survey research is most useful are needs analysis and program evaluation (for more details on what all of these elements are, see Brown, 1989, 1995a). While surveys can be used to

investigate many elements of a curriculum (including objectives, tests, materials, and teaching), it is needs analysis and program evaluation that most often require surveys.

Needs analysis. Needs analysis in language programs is often thought of asthe study of the language forms students will need to use in the target language when they actually try to communicate. The focus is usually on the learners, and their needs are most often stated as linguistic needs. In fact, focusing on the learners in a needs analysis seems reasonable. After all, they are the clients; and the needs of the clients should be met. However, this view may be shortsighted because teachers, administrators, parents, employers, institutions, and even nations may also have needs that should not be overlooked in doing a needs analysis for a particular language program. Needs analysis should probably focus on the language needs of students, but must recognize that students also have other needs, called in Brown (1995a) "situational needs." Situational needs include all of the learners needs as human beings. This means that the definition of needs analysis should be fairly broad. For instance, *needs analysis* might be defined as the systematic collection and analysis of all relevant information necessary to meet the language learning needs of the students within the context of the particular institutions involved in the learning and teaching situation. Since needs analysis relies on input from a variety of different groups about the students' linguistic and situational needs, survey projects may prove particularly useful. In fact, in gathering information from the various groups of people involved, survey interviews and questionnaires may become inevitable. After all, one of the most positive features of survey research is that it can be used to efficiently gather a wide variety of different types of information from a wide assortment of sources.

Program evaluation. Program evaluation might be defined as the systematic collection and analysis of all relevant information necessary to promote the improvement of the curriculum and assess its effectiveness within the context of the particular institutions involved. Since this definition is very similar to the one given above for needs analysis, the reader might correctly conclude that I view the program evaluation process as a kind of ongoing needs assessment. That is true. However, program evaluation is typically based on considerably more and clearer information than needs assessment. While needs analysis is usually done in the early stages of curriculum development and must depend on interview procedures, questionnaires, linguistic analyses, conjecture, and a good deal of professional judgment, program evaluation can take advantage of all of the above information and tools to assess the effectiveness of a program but can also utilize any other information gathered while developing program objectives, developing and analyzing tests, creating materials, and/or teaching the curriculum. *Program evaluation*, then, will be defined here as the ongoing process of information gathering, analysis, and synthesis, the entire purpose of

which is to constantly improve each element of a curriculum on the basis of what is known about all of the other elements, separately as well as collectively. Thus, survey research is a crucial part of program evaluation just as it is for needs analysis and for precisely the same reasons. While survey projects used in curriculum development generally remain in-house, some find their way into print (for examples, see Kimzin & Proctor, 1986; or Richterich & Chancerel, 1978).

Surveys for Language Research

Surveys are also useful for other types of research not directly related to curriculum development. For example, surveys have been used to investigate the characteristics of graduate level teacher-training programs (Ochsner, 1980), or the opinions and attitudes of members of a national professional organization (Brown, Knowles, Murray, Neu, & Violand-Sanchez, 1992). Of all the available research methods, survey research may be the most practical and easy to do because it relies heavily on common sense and less on complex statistics. In fact, survey results are most often reported as simple percentages and averages. Since language survey research can be used to describe, explore, or explain physical characteristics, phenomena, behaviors, attitudes, etc., there is no shortage of useful questions that can be answered by doing survey research. Given that people's ideas, attitudes, opinions, etc. change over time, the possibilities for doing language survey research seem inexhaustible because any research done today might usefully be repeated in five or ten years to see if the participants' ideas, attitudes, opinions, etc. have changed over time.

Steps in Survey Projects

The main problem in setting up a survey project is to make sure that it is planned carefully in a step-by-step manner so the project can be conducted systematically and effectively. Stacey and Moyer (1982) suggest the following ten steps for the construction of surveys:

1. Specifying survey objectives and research questions.
2. Reviewing the literature.
3. Defining abstract concepts.
4. Selecting question formats.
5. Selecting the statistical analysis.
6. Writing the survey questions.
7. Ordering the questions in the survey.
8. Adjusting the physical appearance of the survey.
9. Preparing the cover letter and instructions.
10. Validating the survey.

These ten steps are fine if the only concern is with creating a survey instrument. However, I feel that a language survey instrument should be viewed as part of a larger survey project, one which involves the following steps:

A. Planning the survey project.
B. Designing the survey instrument.
C. Gathering and compiling the survey information.
D. Analyzing the survey statistically.
E. Analyzing the survey logically.
F. Reporting the results of a survey.

If each of these six steps is carefully carried out (along with the ten steps for constructing a survey proposed by Stacey and Moyer (1982) under point B of my list), any survey project will stand a much better chance of success.

Functions of Surveys

Anyone designing a survey instrument must realize that surveys can function in different ways. For instance, surveys can function as: biodata surveys, opinion surveys, self-ratings, judgmental ratings, rankings, or Q-sorts. Obviously, some or all of these functions can be combined in one interview or questionnaire.

Biodata surveys are used to elicit biographical background information about the participants, including things like the participant's age, place of birth, sex, marital status, occupation, number of years of language study, years of education, major field of study, and salary.

Opinion surveys are designed to uncover the opinions and attitudes of the participants about specific issues. For instance, questions might be developed to find out what a group of students think (i.e., their opinions) about the goals and objectives of a particular language program as well as about its tests, materials, teaching, administration, etc.

Self-ratings require the participants to rate their own abilities, interest levels, motivations, etc. Such ratings reveal information about the self-image of the individuals involved.

Judgments can be used to obtain the views of participants about various aspects of language learning. Judgments can be obtained either by asking participants to briefly write their views down or by using a more structured type of question. For example, students might be asked to judge the effectiveness of a language program's objectives and materials in terms of how useful they are, how necessary they are to future language use, or how easy they are to learn. The students would simply be providing their judgments.

Rankings require the participants to order some set of words, or phrases, from most important to least important, or most useful to least useful, etc. For instance, students might be asked to rank a list of the 15 types of classroom

activities used in a particular course from most enjoyable to least enjoyable by numbering them 1 to 15.

Q-sorts combine attributes of one of the earlier survey types with the ranking procedures discussed immediately above. This combination is accomplished by: (a) having the participants give their views, attitudes, opinions, self-ratings, and/or judgments; (b) then asking the participants to rank the issues in terms of how important they think each is. While survey instruments can be designed to perform any one of the six functions discussed above, they are seldom so unidimensional; surveys are most often developed to accomplish a combination of functions.

Guidelines for Writing Good Survey Questions

Once decisions have been made about which functions to use in a survey (or which combinations of functions), the actual questions must be written. Writing good questions can be surprisingly difficult. Guidelines for writing good survey questions are given in Table 2. Overall, the goal in writing survey questions should be to make them so clear and precise that the participants will know exactly what is being asked of them.

Table 2
Things to avoid in writing good survey questions

Avoid:	
1.	Overly long questions
2.	Unclear or ambiguous questions
3.	Negative questions
4.	Incomplete questions
5.	Overlapping choices in questions
6.	Questions across two pages
7.	Double barreled questions
8.	Loaded word questions
9.	Prestige questions
10.	Embarrassing questions
12.	Biased questions
13.	Questions at the wrong level of language
14.	Questions that respondents are incompetent to answer
15.	Assuming that everyone has an answer to all questions
16.	Making respondents answer questions that don't apply.
17.	Using irrelevant questions
18.	Writing superfluous information into questions.

BEST COPY AVAILABLE

Notice that Table 2 is cast in negative terms because it is a list of things that should be avoided in writing survey good survey questions. To begin with, (1) overly long questions, (2) unclear or ambiguous questions, (3) negative questions, and (4) incomplete questions should all be avoided. Overlapping choices (5) should also be avoided. The issue of overlapping choices arises when options are not mutually exclusive. In the following example, the options are overlapping in the sense that Santa Ana High School is located in the city of Santa Ana and in Orange county so the options are not mutually exclusive:

> The best location for the new foreign languages building would be:
> __In Orange County
> __In the city of Santa Ana
> __In Santa Ana High School
> __Other: _____

Questions should also be typed so they are located in one compact place on a single page (6) so that participants will not miss part of the question. Double-barreled questions (7) arise when a question is really asking more than one question. For instance, the following example contains three questions that could be asked separately (one about education spending, one about military spending, and one about language education spending):

> The United States should spend more on education (especially language education) and less on the military.
> Agree_____ Disagree_____ Don't know____

Loaded words (8) are words that suggest an automatic positive or negative response because they are emotionally charged. For instance, in the following example, the word *obviously* would be a loaded word:

> Spanish is obviously the language to study if you live in the Southwestern United States.
> Agree_____ Disagree_____ Don't know____

Leading questions (9), namely, questions that overtly encourage participants to answer in a certain way, should also be avoided. It is also important to avoid questions that may contain issues of prestige (10). For example, when asked, some people will claim that they read more than they do or that they speak a language better than they do because that is the prestigious way to answer. Embarrassing questions (11), or questions that a particular group of participants might find embarrassing (e.g., questions containing swear words) should be avoided, as should questions that show prejudice or bias (12) against a particular group of people.

In addition, questions at the wrong level of language (13) and questions that participants are incompetent to answer (14) should be avoided. And, as matters of sound survey development policy, it is probably best not to (15) assume that

everyone has an answer to all questions, (16) make participants answer questions that don't apply to them, (17) use irrelevant questions, or (18) write superfluous information in questions.

Guidelines for Producing Effective Questionnaires

Writing good questions, as covered in the previous section, is just one of the issues involved in producing sound surveys. Many others are covered in Brown (unpublished ms.)

Table 3
Guidelines for Producing Effective Questionnaires

1. Write good questions (see Table 2)
2. Order the questions rationally
3. Format the questionnaire for clarity
4. Write clear directions
5. Edit carefully

For questionnaire surveys, Table 3 suggests that the ordering of questions is also important. In order to make a questionnaire clear to the participants, it may help to rationally group the questions by their type, function, format, topic, or some combination of these factors. (For more ideas on structuring surveys, see Gallup, 1947).

Formatting the questionnaire for clarity is another important issue, which may involve effective use of spacing, type faces, graphics, highlighting, etc. Writing clear directions is another crucial issue to think about because unclear directions can confuse and discourage participants. Last but not least, it is necessary to use careful editing, perhaps including getting others to edit the questionnaire, piloting it with a few students, and doing very careful final editing.

Guidelines for Increasing Mail Survey Return Rates

There are also a number of ideas that might help in getting a relatively high return rate, especially for those who are using a mail survey, where a questionnaire is mailed out and participants are asked to send it back after filling it out. First, a cover letter explaining the background and importance of the questionnaire may help. Participants are also more likely to return a short questionnaire that is easy to fill out than to return a long and difficult one. If possible, some sort of incentive should be offered to participants for answering the questionnaire (e.g., a monetary reward, or a copy of the study that will result from the questionnaire). Providing return postage and a self-addressed envelope will

also encourage participants to return the questionnaire. It is also a good idea to put the return address somewhere else on the questionnaire so it can be sent back even if the participant loses the envelope.

For participants who do not return the questionnaire, a follow-up letter or phone reminder may encourage them to complete the task. The timing of the mailing may also be significant. For instance, if a questionnaire mailing goes out late in the school year or arrives during the summer break, it is unreasonable to expect academics to respond quickly in great numbers.

Tabulating and Analyzing Survey Results

Once the survey results have been collected, some system must be worked out for tabulating and analyzing all of the data. The system will differ for closed-response questions and open-response questions (both of which may be found on the same questionnaire).

Closed Responses

Closed-response questions are those which offer the participants options from which to select their answers. This category includes Likert scales (the 1 to 5 type), yes/no questions, or any other select-a-response type question. Such questions can be analyzed with straightforward averages or percentages of people selecting each option, or both. For the statistically initiated, standard deviations and chi-squared statistics may prove useful, but the averages and percentages are usually sufficient for finding interesting patterns in closed-response survey results.

In addition, it is often a good idea to check the reliability of a survey instrument—especially if it is being used to make important decisions or to do publishable research. Given that there are often weighted scales on questionnaires, the Cronbach alpha reliability estimate is probably the best to use. Cronbach alpha is available for use on MS-DOS system computers in the SPSS/*PC*+™ (1990) or *SYSTAT*™ (Wilkinson, 1988) software, or it can be calculated by hand using the formula found in Brown (1995b).

If there are subscales involved in the survey, it may be a good idea to examine the reliability of each subscale as well as the combination of all closed-response subscales taken together.

Open Responses

Open-response questions are questions that require the participant to produce an answer, in either an oral or written form. If the responses are oral, they should be taped in the process of gathering the responses. It is usually necessary or at least advisable to transcribe the tapes to a written form.

Even if the open responses are written by the participants, I have usually found it necessary to enter them into computer files for easier analysis. I then do computer analysis of these files using a software package like Right*Writer*™ (Que

Software, 1990) to get frequency counts of the words in the files. Sometimes, as part of the analysis, I use the search function in my word processor to search for the most commonly used words and read how the respondents used them.

With or without the technological boost that computers provide, the open responses should all be read, analyzed, sorted, and summarized in some way. This process may involve simply putting them on cards and getting down on the floor to sort them out, or may involve high technology. Either way, the human brain ultimately ends up doing the analysis and interpreting the results. For examples of interpretations of such open-response questions, see Brown (1992a, 1992c, 1992d, & 1992e).

Suggestions for Further Reading

A number of different topics have been covered in this chapter. Naturally, each of these topics is covered in more detail elsewhere. For instance, more information on the different types, uses, and functions of surveys, as well as ideas for writing good survey questions and getting good return rates can be found in Babbie (1973); Bailey (1982); Brown (unpublished); Oppenheim (1966); or Stacey and Moyer (1982). Examples of how surveys have been used for language curriculum development (especially for needs analysis and program evaluation) are available in Brown (1995a); Kimzin and Proctor (1986); Loschky, Stanley, Cunha, and Singh (1987); Richterich and Chancerel (1978); and Weaver, Pickett, Kiu, and Cook, J. (1987).

Examples of how surveys have been used in language research can be found in Bailey and Brown (1995); Brown (1992a, 1992c, 1992d, 1992e, and 1995a); Brown, Knowles, Murray, Neu, and Violand-Hainer (1991, 1992); Johns (1981); Miller (1995); and Ochsner (1980).

More details on using computers for program level decision making and research can be found in Brown (1992b).

Additional information on the statistics necessary to analyze closed-response survey results, see Brown (1988); Butler (1985); Hatch and Lazaraton (1991); Seliger and Shohamy (1989); and Woods, Fletcher, and Hughes (1986).

More details on increasing the reliability and validity of the open-response analytic processes are available from Agar (1986); Davis (1992); or Kirk and Miller (1986).

Survey interviews and questionnaires provide seemingly simple and straightforward tools that can be used effectively for curriculum development or other types of research. However, language professionals who actually use these tools often find that they are not as simple as they seem. Hopefully, this paper has provided much of the information necessary to effectively plan and conduct surveys, or at least pointed the reader to sources that can help them use surveys responsibly and professionally.

References

Agar, M. H. (1986). *Speaking of ethnography*. Beverly Hills, CA: Sage.

Babbie, E. R. (1973). *Survey research methods*. Belmont, CA: Wadsworth.

Bailey, K. D. (1982). *Methods of social research* (2nd ed.). New York: Free Press.

Bailey, K., & Brown, J. D. (1995). Language testing courses: What are they? In A. Cumming, & R. Berwick (Eds.) *Validation in language testing*. Clevedon, England: Multilingual Matters.

Brown, J. D. (1988). *Understanding research in second language learning: A teacher's guide to statistics and research design*. Cambridge: Cambridge University Press.

Brown, J. D. (1989). Language program evaluation: A synthesis of existing possibilities. In K. Johnson (Ed.) *The second language curriculum* (pp. 222–241). London: Cambridge.

Brown, J. D. (1992a). The biggest problems TESOL members see facing ESL/EFL teachers today. *TESOL Matters, 2* (2), 1,5.

Brown, J. D. (1992b). Using computers in language testing. *Cross Currents, 19* (1), 92–99.

Brown, J. D. (1992c). What is research? *TESOL Matters, 2* (5), 10.

Brown, J. D. (1992d). What roles do members want TESOL to play? *TESOL Matters, 2* (5), 16.

Brown, J. D. (1992e). What research questions interest TESOL members? *TESOL Matters, 2* (6), 20.

Brown, J. D. (1995a). *The elements of language curriculum design: A systematic approach to program development*. Boston: Heinle & Heinle.

Brown, J. D. (1995b). *Testing in language programs*. Englewood Cliffs, NJ: Prentice Hall.

Brown, J. D. (unpublished). *Designing language surveys*. Honolulu, HI: University of Hawaii at Manoa.

Brown, J. D., Knowles, M., Murray, D., Neu, J., & Violand-Hainer, E. (1991). TESOL research questionnaire sent to 1,800 members. *TESOL Matters, 2* (1), 1, 5.

Brown, J. D., Knowles, M., Murray, D., Neu, J., & Violand-Hainer, E. (1992). *TESOL Research Task Force report to the TESOL Executive Board*. Unpublished ms. Alexandria, VA: TESOL.

Butler, C. (1985). *Statistics in linguistics*. Oxford: Blackwell.

Davis, K. A. (1992). Validity and reliability in qualitative research on second language acquisition and teaching: Another researcher comments. *TESOL Quarterly, 26* (3), 605–608.

Gallup, G. (1947). *Qualitative measurement of public opinion: The quintamensional plan of question design*. Princeton, NJ: American Institute of Public Opinion.

Hatch, E., & Lazaraton, A. (1991). *The research manual: Design and statistics for applied linguistics*. Rowley, MA: Newbury House.

Johns, A. M. (1981). Necessary English: A faculty survey. *TESOL Quarterly, 15* (1), 51–57.

Kimzin, G. & Proctor, S. (1986). *An ELI academic listening comprehension needs assessment: Establishing goals, objectives, and microskills*. Honolulu, HI: Unpublished Scholarly Paper, Department of ESL, University of Hawaii at Manoa.

Kirk, J., & Miller, M. L. (1986). *Reliability and validity in qualitative research*. Beverly Hills: Sage.

Loschky, L., Stanley, J., Cunha, C., & Singh, S. (1987). *Evaluation of the University of Hawaii English Language Institute Reading Program*. Honolulu, HI: Unpublished ESL

630 project, Department of ESL, University of Hawaii at Manoa.

Miller, T. (1995). Japanese learners' reactions to communicative English lessons. *JALT Journal, 17* (1), 31–52.

Ochsner, R. (1980). Job–related aspects of the M.A. in TESOL degree. *TESOL Quarterly, 14* (2), 199–207.

Oppenheim, A. N. (1966). *Questionnaire design and attitude measurement*. New York: Basic Books.

Que Software (1990) *RightWriter* ™.

Richterich, R., & Chancerel, J. L. (1978). *Identifying the needs of adults learning a foreign language*. Oxford: Pergamon Press.

Seliger, H. W., & Shohamy, E. (1989). *Second language research methods*. Oxford: Oxford University Press.

SPSS/PC+ ™ (1990). Chicago, IL: SPSS Inc.

Stacey, S. E. & Moyer, K. L. (1982). *A guide to survey development: Manual on writing a survey*. Harrisburg, PA: Pennsylvania State Department of Education. [Also ERIC Document ED 222 538].

Weaver, J., Pickett, A., Kiu, L., & Cook, J. (1987). *Foreign TA training project needs analysis*. Honolulu, HI: Unpublished ESL 630 project, Department of ESL, University of Hawaii at Manoa.

Wilkinson, L. (1988). *SYSTAT™: The system for statistics*. Evanston, IL: SYSTAT, Inc.

Woods, A., Fletcher, P., & Hughes, A. (1986). *Statistics in language studies*. Cambridge: Cambridge University Press.

Chapter 10

Bare Minimum Knowledge for Understanding Statistical Research Studies

Nagata Hiroto

Yokohama National University

This article is for those language teachers who find reading statistical research troublesome, but necessary. First it will examine the statistical reasoning underlying quantitative, empirical studies. Normal distribution, standard deviation, three indices of central tendency (mean, median, and mode), and skewness will be discussed. Then, I will briefly look at the three different kinds of comparisons most frequently reported in quantitative studies, namely, comparisons of means, frequencies, and correlation coefficients. Statistical tools typically used for each comparison will be touched upon. In the course of the discussion, useful checkpoints will be provided that readers of statistical research studies—not to mention researchers themselves—should be equipped with.

Normal Distribution

One of the most interesting phenomena that occurs and recurs in nature, is a pattern called normal distribution (also called bell or normal curve). Suppose, for instance, if you decided to plot on a chart the lengths of a bagful of bean sprouts, the result might look something like the one in Figure 1.1. The shape is not yet "normal," but the larger the number of sprouts you measure, the more "normal" the form of the scatterplot becomes, and it will eventually look some-

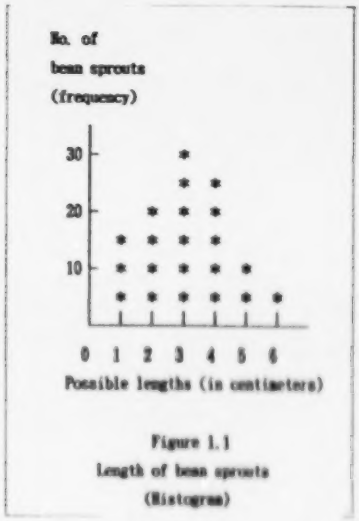


Figure 1.1
Length of bean sprouts
(Histogram)

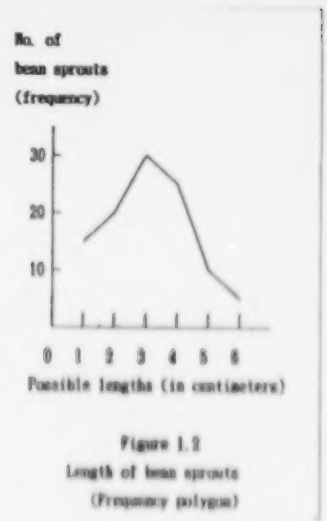


Figure 1.2
Length of bean sprouts
(Frequency polygon)

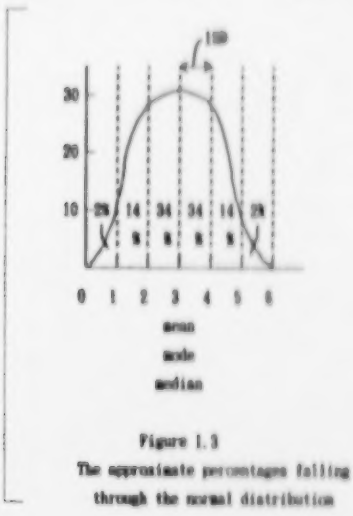


Figure 1.3
The approximate percentages falling
through the normal distribution

thing like the frequency polygon shown in Figure 1.2. As a language teacher, you might want to plot your students' test scores. In this case, most probably, you would also find something close to a normal distribution among your students when the number of them is sufficiently high (usually, at least 50).

Central Tendency

Let us now briefly review three indices of central tendency: the mean, the mode, and the median. The mean is the arithmetic average, and is the most commonly reported indicator of central tendency. The mode is the most frequent score in a *set of data*. The median is the score which divides the entire data in half. If the mean and the median coincide, and if such a distribution has a single mode, the frequency distribution will show an ideal symmetrical curve.

What is intriguing about such normal distribution is that we would expect about 68% of all the scores to fall within one standard deviation of the mean, and 95% to fall within two standard deviations (see Figure 1.3 above). Standard deviation shows dispersion, or in other words, how much the scores vary away from, or spread out around, the mean (Brown, 1991, p. 574).

Figure 2.1 A Symmetrical Distribution

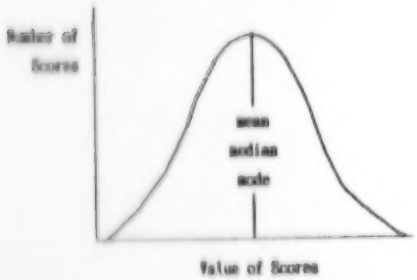


Figure 2.2 A Negatively Skewed Distribution

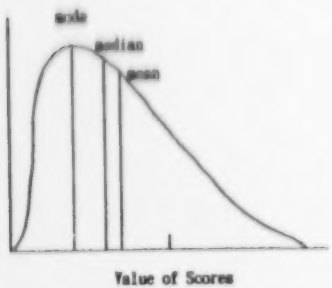
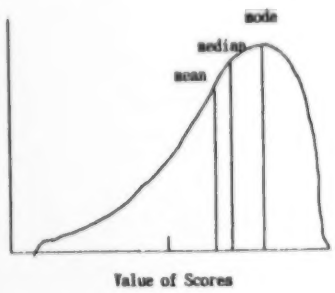


Figure 2.3 A Positively Skewed Distribution



BEST COPY AVAILABLE

Table 1 Descriptive Statistics for Comprehension

Group	m	SD	min	max	n
G1	3.80	2.20	.0	10.0	75
G2	2.76	1.67	.0	9.0	75
G3	4.79	2.47	.0	13.0	75
G4	4.04	1.88	.0	10.0	75

This normal distribution, in its perfect form, however, rarely occurs in the field of language teaching unless the number of subjects is extremely large. The values of the measures of central tendency usually differ. When the value of mean is influenced by the size of extreme scores, it is pulled toward either end of the distribution in which the extreme scores lie, as shown in Figures 2.2 and 2.3, and the distribution will show an asymmetrical curve (described as being skewed).

When we read statistical studies, we should look to see how far the distribution is skewed from the normal curve, because statistical analyses are based on normal distribution. If you find the distribution is extremely skewed (not close to normal), you should look for the kind of interpretation that is made by the researcher of the study. The skew of a distribution can be identified by comparing the mean and the median without necessarily constructing a histogram or frequency polygon (see Figures 1.1. and 1.2 for examples). When the distribution is skewed toward the lower end, or negatively skewed, the mean is always smaller than the median, and the median is usually smaller than the mode (see Figure 2.2). When a distribution is skewed toward the higher end, or positively skewed, the mean is always greater than the median and the median is usually greater than the mode (see Figure 2.3). Unfortunately, most statistical studies do not provide information on the median. Thus, we usually have to turn to other indices usually provided in the table called "descriptive statistics." Table 1 is an abbreviated version of one such example. It displays such vital information as the number of students involved in the study (n), the mean (m), the standard deviation (SD), and minimum (min) and maximum (max) scores.

As is shown, four groups (G1 to G4) each consisting of 75 subjects took a comprehension test of some kind, and their scores ranged from zero to 13. The means (m) for all groups are pulled slightly towards the lower ends. We know this because all mean scores are lower than the halfway point scores between the lowest possible scores of zero and the highest possible scores of ten, nine,

thirteen, and ten for each group, respectively. However, the scores of each group are spread out to a reasonable degree (we know this because there is room for two standard deviations above and below the mean within the range of the lowest possible and highest scores), therefore, we can conclude that the skewness is not a serious problem here. The table also shows that in terms of means, Group 3 comes first, and Groups 4, 1, and 2 follow. The standard deviation (SD) for Groups 1 and 3 are higher, however, indicating more spread or greater dispersion.

What we would like to know, then, is whether these mean differences are significant. Can we confidently say that they are really (statistically) different? In other words, did the differences we observed happen accidentally, or did they manifest themselves because of some other systematic factors? In order to answer this question, researchers conduct statistical analyses. (For a clear and concise account on the logic of statistical inference behind these analyses, see Nunan, 1992, pp. 28–37).

Three Different Kinds of Comparisons

The three most common types of statistical tests typically reported are mean comparisons, frequency comparisons, and comparisons of correlation coefficients to zero. Let us look briefly at each of them.

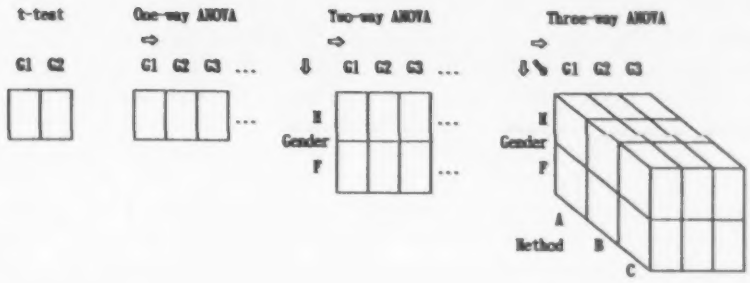
Mean Comparisons

A mean comparison is comparing two or more groups by comparing their average scores on some test. There are four kinds of statistical tests typically used within this category. They are t-test, one-way ANOVA, n-way ANOVA (factorial design), and MANOVA.

The t-test is used to compare the means of two groups, whereas one-way ANOVA is used to examine the differences in more than two groups. N-way ANOVA enables us to analyze the effect of different treatments in more complex conditions, such as different proficiency levels, or different types of learners (e.g., learners with different learning styles). Therefore, the comparisons are drawn in more than two directions (actually, N directions) in n-way ANOVA. In other words, there are more than two (N) independent variables. However, ANOVA is just an exploratory tool, which only tells the researcher that there are some significant differences somewhere in the data but does not specify where they are. Subsequent analyses are needed for the researcher to identify where the significant differences lie (see the one-way ANOVA section below). Still, however, no information can be obtained about the magnitude of the effect.

Diagramatic representations of each of the tests are on the following page.

Now, let us use the data in Table 1 and actually run these statistics to see if we can make any sense out of the computer printouts.



Diagramatic representations of types of designs appropriate for t-test, one-way ANOVA, and n-way ANOVA

T-test: For mean comparison between two groups. For the sake of simplicity, let us assume there are only two randomly chosen groups (groups 2 and 3 in Table 1) learning a set of new vocabulary, and we are comparing the performances of these two groups. Group 2 learned a set of vocabulary using only a listening task, whereas Group 3 learned the same set through the same listening task, but with the help of written input. After the task, we investigate the effect of the written input by giving both groups the same vocabulary test.

Below is a computer printout (SPSS/PC+) of the t-test results. It shows the t-value to be -5.89, which is significant at the .05 level. This means that Group 3 subjects who learned a set of vocabulary through a listening task with the help of written input learned better than those in Group 2. This will be reported as: $t = -5.89$; $p \leq .05$. This significance level (called alpha decision level) should be set at the beginning of the study. Usually, it is set at the conservative $\alpha < .01$, or at the more liberal $\alpha < .05$.

One-way ANOVA: For comparing differences in more than two groups. Suppose there were four groups (groups 1, 2, 3, and 4 in Table 1). Each group was given a different kind of treatment, for instance, they were taught by using the G-T method, TPR, Silent Way, and Suggestopedia. Now, we want to know if these different treatments had any effect. Since the comparisons are made in only one direction, that of the independent variable, which is "method," this ANOVA is called a "One-way" ANOVA.

Below is a computer printout of the one-way ANOVA. F probability (.0000) indicates that there were significant differences between the four groups, but it

Independent samples of GRP

Group 1: GRP EQ 2.0 **Group 2: GRP EQ 3.0**

t-test for: COMP

	Number of Cases	Mean	Standard Deviation	Standard Error
Group 1	75	2.7600	1.667	.193
Group 2	75	4.7867	2.468	.285

		Pooled Variance Estimate			Separate Variance Estimate		
F Value	2-Tail Prob.	t Value	Degrees of Freedom	2-Tail Prob.	t Value	Degrees of Freedom	2-Tail Prob.
2.19	.001	-5.89	148	.000	-5.89	129.91	.000

[SPSS/PC+ printout 1]

does not specify where those differences lie. Also, although descriptive statistics (mean and standard deviation) provides an indication of where the differences are likely to exist (e.g., of all the mean differences, that between G1 and G4 is the smallest), these insights must be checked statistically. The post hoc Scheffe contrast test does this job. As shown at the bottom of the printout, where an asterisk (*) denotes pairs of groups significantly different at the .05 level, differences exist between Groups 1 and 2, 2 and 3, 2 and 4, and 1 and 3, but not 1 and 4.

Readers are advised not to use multiple t-tests here. Although there have been several studies using multiple t-tests published even in prestigious journals in our field, using t-tests repeatedly (for multiple comparisons) is not recommended because the more comparisons you make, the more chances of creating spurious results.

N-way ANOVA: For examining the effect of several variables studied simultaneously, as well as the interactions among the variables. As is shown diagramatically above, the comparisons are made in more than two directions (i.e., in N ways) in N-way ANOVA. (For a more detailed account, see Hatch & Lazaraton, 1991, pp. 301–331.)

MANOVA: For research designs including more than one dependent variable. Language studies often include two or more dependent variables which are related to each other. For example, if three different kinds of tests were given to

---------- ONEWAY ----------

Variable COMP
By Variable GRP

Analysis of Variance

Source	D.F.	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	3	157.8000	52.6000	12.1909	.0000
Within Groups	296	1277.1467	4.3147		
Total	299	1434.9467			

---------- ONEWAY ----------

Variable COMP
By Variable GRP

Multiple Range Test

Scheffe Procedure
Ranges for the .050 level -
3.98 3.98 3.98
The ranges above are table ranges.
The value actually compared with Mean(J) - Mean (I) is ..
1.4668 * Range * Sqrt (1/N(I) + 1/N(J))

(*) Denotes pairs of groups significantly different at the .050 level

---

---------- ONEWAY ----------

Variable COMP
(Continued)

Mean	Group	Grp 2	Grp 1	Grp 4	Grp 3
2.7600	Grp 2				
3.8000	Grp 1	*			
4.0400	Grp 4	*			
4.7867	Grp 3	*	*		

[SPSS/PC+ printout 2]

LIFE Is life exciting or dull? by SEX Respondents's sex

SEX Page 1 of 1

Count / Exp Val / Residual	Male 1	Female 2	Row Total
LIFE			
1 Excited	300 279.0 21.0	384 405.0 -21.0	684 46.8%
2 Not excited	296 317.0 -21.0	481 460.0 21.0	777 53.2%
Column Total	596 40.8%	865 59.2%	1461 100.0%

Chi-Square	Value	DF	Significance
Pearson	5.00467	1	.02528
Continuity Correction	4.76884	1	.02898
Likelihood Ratio	5.00327	1	.02530
Mantel-Haenszel test for linear association	5.00124	1	.02533

Minimum Expected Frequency = 279.031

Number of Missing Observations: 12

[Chi-square printout]

the same subjects in two different groups, the statistical test called for is not ANOVA, but MANOVA, because the scores from the three different kinds of tests administered to the same subjects should somehow be related. These

Correlations:	FIN	MLAT5	PLAB5	PLAB6	MLAT4	COMP
FIN						
MLAT5	.1735*					
PLAB5	.1233	.1346*				
PLAB6	-.0231	.0796	.2317**			
MLAT4	.0326	.2718**	.1071	.2111**		
COMP	.0033	.3876**	.1032	.0083	.2126**	
PT1	.1424*	.2582**	.1615*	-.0656	.1241	.4303**
PT2	.0656	.2191**	.0791	.0250	.1188	.4474**
PT3	.1100	.1199	.1047	-.0738	.1787**	.3083**

N of cases: 300 1-tailed Signif: * - .01 ** - .001

Correlations:	PT1	PT2	PT3
FIN			
MLAT5			
PLAB5			
PLAB6			
MLAT4			
COMP			
PT1			
PT2	.6597**		
PT3	.6137**	.6634**	

N of cases: 300 1-tailed Signif: * - .01 ** - .001

[SPSS/PC+ printout 3]

related variables should be analyzed together, not separately. (For further details, refer to Hatch & Lazaraton, 1991, pp. 386–387).

Frequency Comparisons

Chi-square: For comparing two nominal (frequency) data. Frequency means counting the number of times something happens. When examining relations between frequencies, the Chi-square analysis is the usual procedure employed. For example, suppose you conducted a survey asking people whether or not they think life is exciting. One thousand four hundred and sixty one persons

responded, and 50.3% of all the men (300 out of 596) and 44.4% of all the women (384 out of 865) described life as exciting. Is this evidence sufficient enough to believe that men differ from women in finding life exciting? To answer a question like this, the Chi-square test is calculated.

As shown in the computer printout, this procedure compares the observed frequencies and expected frequencies to see if the former are greater than chance alone. The observed significance level of a Chi-square of 5.005 is 0.025. This indicates that a discrepancy this large between the observed and expected frequencies would occur only 2.5% of the time if, in the population, men and women are equally excited with their lives. Since the observed significance level is quite small, we can conclude that men and women are not equally likely to find life exciting.

One important fact we should bear in mind here is that this family of analyses should be used to research where only nominal variables are included and frequencies are compared, and that, more important, the Chi-square procedure does not allow us to make cause–effect claims.

Comparisons of Relationship

Pearson correlation: For examining existing relationships between variables without any manipulation of the variables (e.g., no treatment). For example, if we wonder whether a good language aptitude is related to success in vocabulary learning and acquisition, we might want to know the relationship of the subjects' scores on aptitude tests as well as their scores on a vocabulary comprehension and retention test. Below is the computer printout of one such study (Nagata & Ellis, 1996, Forthcoming). FIN, MLAT4, PLAB5, PLAB6, and MLAT4 are subtests of aptitude batteries. COMP, PT1, PT2, and PT3 are vocabulary comprehension and retention tests. The asterisk (*) shows where significant relationships were detected.

Here again, as with the Chi-square analysis, we should be cautioned that the correlation tells us only that there is some degree of relationship between the two variables. We cannot, therefore, make any cause–effect claims.

Statistical Assumptions to Remember

Selecting an appropriate statistical test is a crucial point in any statistical study. Major questions we should ask ourselves are: (a) How many variables are there and what are their functions in the study (independent vs. dependent variables)?; (b) What types of measurement tools (scales) were used?, and; (c) Where do the data come from, i.e., from two different groups (a between-groups design), or two or more measures taken from the same group (a repeated-measures design)?

Even after an appropriate test has been selected, we should not forget the fact that all statistical tests have certain assumptions underlying their formulation. When

even one of these assumptions is violated, the probability gets distorted, and it becomes difficult to know how much confidence to place in the findings. Check to see if these assumptions are met. If you are the reader of a statistical study, see if the researcher checked whether the basic assumptions of the specific test(s) employed were met. (For a compact list of assumptions and solutions when these assumptions cannot be met, see Hatch & Lazaraton, 1991, pp. 546–554.)

By Way of Conclusion

I have taken you quickly through the winding mountain path of basic statistical manipulations often used in our field. Just to recapitulate some of the checkpoints which have so often been neglected in the statistical studies published in our field:

1. Check to see if the appropriate statistical tests are used.
2. Check to see if all the statistical assumptions are met.
3. See if the number of subjects is sufficient. (Rule of thumb says more than 30.)
4. Check to see if multiple comparisons are being made erroneously. Especially watch out for the use of the ordinary t-test for making multiple comparisons).
5. See if the nominal, frequency data are properly dealt with, in other words, by the Chi-square analyses for independent data.
6. Check to see if the researcher's interpretation of the findings stays within the statistical logic. For example, the Chi-square procedure is a non-parametric test. It does not allow us to make cause–effect claims. Neither does the Pearson correlation.
7. Check to see if the dependent variables in the study are related. If so, use MANOVA, instead of ANOVA.
8. See if the data are from two different groups or are two or more measures taken from the same group. In other words, be careful about the repeated-measures designs. Data, then, should be examined in terms of within-subjects differences rather than between-subjects differences. If, for instance, a pretest and a posttest are administered, the study is repeated-measures in design.

References

I would like to use this section not only as a conventional reference section, but an introductory resource section. Therefore, the books and articles are not in conventional alphabetical order. The following are books and articles which I have found accessible, understandable, and extremely readable. If you are new to statistics you might like to start reading in the order they are presented here.

1. Brown, J. D. (1988). *Understanding research in second language learning*. Cambridge: Cambridge University Press.
2. Seliger, H. W., & Shohamy, E. (1989). *Second language research methods*. Especially, Chapter 9: Analyzing the Data (pp. 201–242). Oxford: Oxford University Press.
3. Brown, J. D. (1991). Statistics as a foreign language—Part 1: What to look for in reading statistical language studies. *TESOL Quarterly, 25* (4), 569–586.
4. Nunan, D. (1992). *Research methods in language learning*. Especially, Chapter 2: The Experimental Method (pp. 24–51). Cambridge: Cambridge University Press.
5. Hatch, E., & Lazaraton, A. (1991). *The research manual: Design and statistics for applied Linguistic*. Rowley, MA: Newbury House.
6. Brown, J. D. (1992). Statistics as a foreign language—Part 2: More things to consider in reading statistical language studies. *TESOL Quarterly, 26* (4), 629–634.

Chapter 11

In-Depth Interviewing as Qualitative Investigation

Marilyn Books

Waseda University

> Interviewing is rather like marriage: everybody knows what it is; an awful lot of people do it, and yet behind each closed door there is a world of secrets (Oakley, 1986, p. 231).

Social behavior is complex and multidetermined, and there is no perfect method for studying it. Of the variety of methodological approaches, this paper presents in-depth interviewing as a strategy in qualitative research for discovering the secrets of the complex human psyche. Krathwohl (1993, xvii) affirmed that most research texts:

> hew to a positivistic point of view, stressing experimentation and ignoring or minimizing qualitative research . . . . It is clear that we have entered an era where more than the positivistic paradigm is considered defensible.

What determines the use of a qualitative method such as in-depth interviewing?: the purpose of the study does. If the question is, "How does the degree of adoption of the American custom of explicitness correlate with the length of overseas experience?" then a survey may be the best choice. If the scholar is wondering, "How do Japanese professors speak with their American cohorts in staff meetings?" then participant observation would be a likely technique. If the objective is to determine whether Japanese language lessons hasten the acculturation of foreigners, then an experimental method may suit best; however, if the purpose is to

BEST COPY AVAILABLE

comprehend the meaning of people's existence ("What is it like talking to a Japanese person who is speaking indirectly? What is your experience with that/them?") then open-ended oral questioning is a useful avenue of study.

This paper begins by looking at the nature of in-depth interviewing, situating it on the continuum of research methodology, and differentiating it from quantitative research. It then discusses the strengths and weaknesses of open-ended surveying. Sampling to determine interviewees will be discussed in the next section. Finally, some advice on the actual administration of the interview is offered.

The Sample Study

The sample study discussed in this paper was originally the basis of a doctoral dissertation in communication (Books, 1996). The topical question was: "What are the positive and negative experiences of North American native English speakers and native Japanese speakers who teach English at Japanese universities in relation to these key communication styles: Americans' assertiveness, explicitness, and independence; and Japanese indirectness, *nemawashi* (groundwork), and *tatemae* (stated position)?"

Research Methods

Research methods fall along a continuum between quantitative and qualitative approaches. Quantitative methodologies generally use numeric data to portray phenomenon, seek to validate hypotheses, preplan with variables and concrete notions about how to explain occurrences, maintain controls (sometimes extensive), employ structures (often highly constrictive), and are often controlled in a laboratory setting.

Qualitative methods usually employ verbal data, begin with few or no preconceived notions, seek to comprehend and explain or describe phenomena, maintain little or no control and structure, explore with the aim of discovering what exists that is of importance in the situation, engage in a more natural function, and are carried out in natural environments.

Some methods span the continuum and can be, depending on the situation, either quantitative or qualitative. Methods which are generally quantitative and stress validation are: experimentation (complete with dependent and independent variables), and model building and testing (using relationships as predictors). Further away from the quantitative pole are: meta-analysis (studying similar studies) and longitudinal studies (long-term studies of groups). Those which typically lean more to description and rely at least somewhat on explanation are: single-subject studies (long-term studies of individuals), evaluation methods (determining efficacy), historical methods (examining personal and institutional documents), surveys (usually closed questionnaires), and classification

schemes (placing material into useful categories). Theory building (exceeding actual data) is mostly description and explanation.

At the other end of the continuum is qualitative research. It must be noted that action research is not really a method, but a useful approach for using any of the available research methods, and such studies are carried out in tandem with the benefactor, for example, the educator who can use the results. A more thorough discussion of the variety of research methods can be found in Krathwohl (1993).

Subsumed under qualitative research are these fundamental strategies: observation, participant observation, document reviews, and interviews. Observation involves noting and recording behaviors, events, and objects in the setting studied. Participant observation immerses the researcher in the setting and experiences of those observed. Included in document reviewing is the perusal of archival data.

Patton (1990) noted that interviews vary in the determination and standardization of questions *before* the interview. He divided open-ended interviews into three categories: the informal conversational interview, the standardized open-ended interview, and the general interview guide approach. The conversational style entails a spontaneous flow of questions. The standardized interview employs a specific list of questions to be asked sequentially. The general interview uses an outline of issues, but the order of presentation may vary and the actual wording is not standardized. It is free-flowing and is regarded as "a speech event" in the words of Mishler (1986, p. 35). That was the style chosen for the study referred to in this work. For example, the American participants were shown a chart with the headings Postives and Negatives; and the terms "indirectness," "*nemawashi* (groundwork)," and "*tatemae* (stated position)." They were then asked, "What are your positive experiences with Japanese indirectness?" or "What are the good points related to the indirectness of your counterparts?" or a similarly worded question.

Differentiating Quantitative and Qualitative Research

I will now compare in-depth interviewing specifically with quantitative research in general.

Apart from a different emphasis on numbers and words, the greatest contrast between quantitative and qualitative investigation is the amount of control the investigator has over the research. In qualitative interviews, interviewees frequently take the topic and develop it, providing new direction and allowing the interview to change course and the interviewer to explore unexpected topics.

Quantitative methods begin with a hypothesis and use a deductive process, whereas interviews gather data and inductively process it to an explanation. Instead of validating an explanation, interviews develop an explanation, often uncovering unexpected results. In the sample study, for instance, the Japanese participants' opinions on American assertiveness were not what the interviewer had expected when the investigation was begun.

Another function of the interview is to reveal a relationship between the differing views of the interviewee and the interviewer (Krathwohl, 1993). As the interviewer in the sample study, I felt the speech strategy of explicitness to be preferable to ambiguity because it fosters comprehension of the speaker's stand. However, it would have defeated the purpose of the interview to argue with the Japanese interviewees who reacted negatively to explicitness:

- "If they [Americans] give too much background information that's boring and insulting—when they confront you with an issue and use a lot of words to be very definite and you know all about it. But it's a habit of Americans."
- "For you the verbal code is supreme; you want everything in writing, like your teaching contract. It's almost as if you can't trust anyone."
- "I find him patronizing and boorish but I guess you find him eloquent."

Unlike quantitative research, in qualitative research most of the creative work—coding and analyzing the revelations into useful slots and assigning meaning to them—follows the data collection.

The objectives of positivistic research and interview research are diametrically opposed to each other. The former deals with validation, while the latter—employing questions like, "Do you like or dislike the assertiveness of Americans? Can you describe your experience with that?"—has basically two goals: description and explanation. In the social sciences in particular, the goal of research is seen as exploring for "the development of explanatory concepts that help you understand individual behavior and social processes, concepts that help make life intelligible" (Jones, 1996, p. 7). To explain a preferred or abhorred communication style is to give a reason for it—why people feel the way they do. In the sample study, reasons were articulated by a middle-aged female Japanese instructor who had lived and taught in Yarmouth, Maine for six years as she explained why she had come to accept assertiveness as positive.

> Well, I've never seen or met an American professor who is not assertive. I like it very much. I mean it's very productive. I don't have to guess to get their point. They come out with a point; then I can be assertive myself, right? Assertive persons usually permit the other person to be assertive, too, and I enjoy the interaction.

A communication scholar who studied in the US for several years and deals with Americans often in his professional associations clarified:

> They assert their own opinion, I think, rather than evading the issue and I appreciate that fact. I take it positively. At a conference, I'd read a paper and one of the Americans came to me after the presentation and told me that he didn't agree. You know, we had a very fruitful exchange of opinions and afterwards we became friends. That's very difficult among Japanese. That was a positive experi-

ence. And during the speech Americans challenged me. Sometimes that shows a reaction to what I'm talking about and I enjoy it. It shows someone's listening to me, not sleeping.

The Strengths of In-depth Interviewing

Perhaps one of the most advantageous aspects of interviews is the fact that they are particularly successful in obtaining large amounts of data in relatively short periods of time.

Open-ended interviewing is typically as close to a natural process as possible. There is more freedom in the type of questions, in the wording of the questions, and in the choice of setting all of which adds to the naturalness. This in turn facilitates cooperation.

In addition, the range of data is unlimited. Experiments begin with a hypothesis and seek to verify it whereas in-depth interviewers are not so restricted. This allows the researcher to explore and discover the topic jointly with the interviewees because of the unstructured way in which the participants express their own thoughts, with a reduced likelihood of missing new and unreported details or concepts. In the sample study, for example, discoveries emerged from the data as I listened to Japanese opinions on American assertiveness:

- "They have warmth because they come up to me and tell me what they think. I like that."
- "I see they go right up to some other Japanese teachers and state their case. They exude confidence. It's fine, I think. . . ."
- "It sure beats the antithesis, being self-effacing, which is what my colleagues [Japanese] try to do. That's [the habit of] a lot of Japanese, and that's not good."

These were not explanations I had thought of when I entered the inquiry. The wholehearted acceptance and preference of assertiveness was contrary to my expectations.

Because the study focused on lengthy individual participation, I was able to better grasp the complexity and background of the individual's experience. This was aided by face-to-face contact in which non-verbal behavior could be observed and noted. Moreover, the peculiar format of interviewing allowed immediate follow-up questioning for expansion, omissions, clarification, and contextualization. These latter features point to how interviewers affect the interview positively: even if unintentionally they give feedback via body language or prompts, encouraging more complete responses.

Rich verbal descriptions which portray phenomena are characteristic as in these responses to the concept of *nemawashi* (groundwork):

> Well, at least *nemawashi* saves the struggle. In Japan there's no struggle, it's smo-o-o-oth. If there were any arguing, it's done privately. Back home they argue vociferously in public, at the meetings, and I used to like the challenge of the confrontation. There's always two camps, the proponents and the opponents. We have a vote and the winners savor the victory, but the losers cause a disagreeable gap and grudgingly wait and plan viciously for the next vote. In Japan that dissension is avoided. I'm turning Japanese. [He laughed.]

> I think *nemawashi* worked, for example, when my visa problems were sorted out. All six of the foreigners consulted each other on my behalf and then we got together and went *en masse* to the department head who in turn went to the personnel in charge of external affairs, who went to his superior, etc. It went all the way up and was then implemented *toute de suite*.

Readers of the published results also frequently identify with the quotations which have intuitive appeal for them.

The Weaknesses of In-depth Interviewing

The gravest weakness of interviewing is the amount of time it takes, not just for the interviewing itself, but also for setting up appointments, commuting to meet the participants, transcribing the tape recordings, and analyzing the data.

Since a one hour recording can take a typist from two to four hours (depending on the accent, grammar, articulation, rate of speech production, and so on) to type, all data in the sample study was transcribed by hired transcriptionists. Additionally, managing the mountains of material typed up presented huge problems. If a response comes at an average rate of 150 words a minute, a one-hour cassette can yield 9,000 words, or 36 pages of text (at 250 words per page). It can be an overwhelming task to sift through, dissect, code, and analyze.

The presence of the recording mechanism can be a deterrent to free speech. In this research example it was not, given the cooperativeness of the participants and the frankness of their statements, for example:

- "So you thought we are all dumb Japanese?"
- "They illiterate [*sic*], on and on and on like a drunkard."
- "You maybe look and act like *ronin* (leaderless samurai).

Another weakness is that errors of commission can occur on the part of the interviewee and the interviewer. There is no assurance that what is related by an interviewee is accurate. The interviewee may also be uncomfortable or fearful and therefore unable or unwilling to speak frankly, and there is no guarantee that the interviewer will not (knowingly or unknowingly) influence the responses.

Errors of omission can also occur. People are commonly unaware of—or unable to verbalize—circumstances in their lives, or may withhold incriminating

evidence. Failure on the part of the interviewer to question some participants in a way that is meaningful to them may also yield parsimonious replies. The role of interviewer demands well developed listening skills and interpersonal communication skills which, if not present, can easily and seriously weaken the quality of data collected.

Analyzing the data presents a whole set of pitfalls. Only two are cited here: (a) How do you explain unanticipated results?, and (b) Verbal data, even more than statistics, is open to various interpretations.

Sampling of Interviewees

> It is always nice to find a way of doing things that will save you time and energy, provided that you do not sacrifice quality in the process. . . . Sampling is just such a time and energy saver. Furthermore, when done properly, it not only does not sacrifice quality, but may actually improve the quality of the data obtained (Jones, 1996, p. 180).

For in-depth interviewing, the interviewer could choose a true representative sample (such as a random, stratified, or systematic sample) of the population of interest as if quantitative research were to be carried out, or vary that by engaging in non-probability sampling such as in the judgmental and purposive format.

For the communication style investigation carried out in the sample study, the "purposive" sampling format was chosen. It is also called "judgmental," "deliberate," or "selective" because the researcher uses judgment in selecting individuals who will be instrumental in gathering data (Lonner & Berry, 1986, p. 87). Patton also uses the term "purposeful," and offers the rationalization that by selecting cases for study in depth, "the logic and power" is revealed: one can learn a great deal about issues of central importance to the purpose of the study, thus the term "purposeful sampling." He contrasted quantitative and qualitative investigation on the basis of sample size and method:

> Perhaps nothing better captures the difference between quantitative and qualitative methods than the different logics that undergird sampling approaches. Qualitative inquiry typically focuses in depth on relatively small samples, even single cases (n = 1), selected purposefully. Quantitative methods typically depend on larger samples selected randomly (1990, p. 169).

For the sample study, the purposive method was chosen instead of random sampling to increase the depth of data exposed and to illuminate the questions under study. In addition, "random or representative sampling is likely to suppress more deviant cases . . . as well as the likelihood that the full array of multiple realities [may be] uncovered" (Lincoln & Guba, 1985, p. 40). An even greater danger of random sampling, however, is the opposite situation, namely, a profusion of extreme cases.

Purposive sampling involves choosing just those elements (each individual in the population is considered an element, or a unit of analysis) you want to interview without regard to generalizability. The selection of participants could be guided (as it was in the cited work) by the premise that the knowledge of self and the ability to communicate provides a rich base of data and has priority over the representativeness of the entire sample. That means that the investigator could plan not to work with great numbers of deviants, that is, the polar cases which would skew the results, provide less than useful material, or not be applicable to the population who would value the results of your study. In effect, purposive sampling can lead to a better representation of the intended consumer.

The strategy in the sample research was the following:

1. I devised a sample frame (a tentative list) of all possible types of candidates from amongst North American native English speakers and native Japanese speakers who teach English at Japanese universities. The scope was delimited to "nativeness," nationality, subject taught and occupation, institution, level of institution, and geographical location: native speakers from Japan and North America teaching English as professors at Japanese colleges and universities in Tokyo.

2. From a basic sample, I reserved a list of aberrant entries such as young, new arrivals, potentially lacking in self-awareness who are often at the "Honeymoon Stage" of the culture shock curve (loving every novel aspect about Japan); very old professors, particularly those about to retire; and those who claimed to be totally bilingual and bicultural. Interviewees had to meet certain other specifications. One general specification was that Japanese subjects had to speak English well enough to discuss abstract topics in depth.

3. In addition, I wanted the sample to take into account the disproportionate number of males hired in Japanese universities. I wanted to make sure that females were included in appropriate numbers corresponding to their presence. I treated each gender as a separate population, and took the same percentage from each, thereby including more males than females.

4. I sought a mix of participants who I knew well and those who were strangers. I felt that friends and acquaintances would provide context, facilitating analysis; while those with whom I had not yet established a rapport might be less hesitant to provide stronger statements and more divergent positions (this proved true).

5. After interviewing my core sample and formulating some generalizations, I tested the robustness of those generalizations via triangulation. I resurrected the aberrant list and selected subjects who might contradict my findings.

Only a small portion of the number of potential candidates were chosen. The sample group consisted of 25 individuals of each racial group, totalling 50 interviewees.

Administrative Techniques

Both science and art are involved in the interview itself. The participant's motivation, the manner of the researcher, and the setting are amongst the most crucial features affecting the quality and validity of the data received, and there is much that the interviewer can do to improve them.

The participant can be motivated by skillful prompting. Utterances such as "And?" and "Then?" and "Mm" can aid the process. Non-verbal communication can be employed like nodding, gestures of encouragement, and moving in your seat as if to get a better (ear) vantage point. Exploring laughter and tolerating silence are among the many good hints provided by Seidman (1991). Douglas stated that:

> In general, the more of a listener the interviewer is, the more sensitive, the more openly intimate, the more sincerely interested in and warmly caring about the other person, the more he can temporarily surrender to the experience and soul of the other person, the more successful he will be as a creative interviewer (1985, p. 57).

It seems obvious that the interviewer should be a good listener, but too often the researcher fails to take in the interviewee's points. This becomes clear when the tapes are played back and missed opportunities to clarify stances or to prod the speaker are noticed. Interviewers should be masters at forgetting their frames of reference, attending to the interviewees, coming to the meeting prepared to be surprised and to learn something entirely new, and entering the interview anxious to have their preconceived notions overturned. The axiom of Heraclitis applies, "If you expect not the unexpected, you shall not find truth." One of the most important attitudes for the interviewer to adopt is that the interviewee's contributions are valuable and useful.

Even with the guide type of interview, hypothetical examples have to be prepared to elucidate concepts which may not be understood. In the sample study, for example, examples of *tatemae* (stated position) were provided when necessary, in the cases where informants were unfamiliar with the term, or when contrasts were made.

As the interviewer will not know many of the interviewees, it is probably best to let the participants choose the setting, with the prompt that "situational encapsulation" (Douglas, 1985, p. 99) is important to avoid interruptions and to foster reflection. Douglas also counseled that, "the recorder is both a reassurance of the seriousness of your pursuit and a brutal technological reminder of the human separateness that undermines the intimate communion you are trying to create" (1985, p. 83). In the sample study, a recorder was used in each interview and was found to be indispensable. More speaker autonomy can be attained if the interviewees are in charge of the pause button allowing them to choose which statements are recorded.

There are certain research variables which must be controlled in both quantitative and qualitative investigation. Brown's quantitative work (1988, pp. 33-35) mentions a few. The halo effect is "the tendency among human beings to respond positively to a person they like"(Brown, 1988, p. 33). Subject expectancy and researcher expectancy figure in both. The reactivity effect is exemplified by the situation in which the "subjects actually form or solidify attitudes that they did not have before [the research]" (Brown, 1988, p. 35).

Other useful guidelines are offered by Frey and Oishi (1995) on logistics (availability of resources such as facilities) and data quality (e.g., confidentiality).

Conclusion

The underlying cultural or social motivations for a behavior, communicative or otherwise, are central and complex, and that is what makes investigation of human behavior both difficult and exciting. Which method to pursue is the challenge. There is no perfect method.

Qualitative methodology—in particular interviews—have been overlooked for too long. The compelling strengths of inductive, intensive, individual, insightful interviews will propel them to greater popularity and a higher status in education and other social science fields in the twenty-first century as they make valuable contributions in uncovering "the world of secrets" out there.

References

Books, M. (1996). *Communication styles of Japanese and Americans: Six key styles explored*. Ann Arbor, MI: University Microfilms Inc. (University Microfilms No. 95-42, 658.)

Brown, J. D. (1988). *Understanding research in second language learning: A teacher's guide to statistics and research design*. New York: Cambridge University Press.

Douglas, J. D. (1985). *Creative interviewing*. Beverly Hills: Sage.

Frey, J. H., & Oishi, S. M. (1995). *How to conduct interviews by telephone and in person*. Thousand Oaks: Sage.

Jones, R. A. (1996). *Research methods in the social and behavioral sciences* (2nd ed.). Sunderland, MA: Sinauer Associates.

Krathwohl, D. R. (1993). *Methods of educational and social science research: An integrated approach*. New York: Longman.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Beverly Hills: Sage.

Lonner, W. J., & Berry, J. W. (1986). Sampling and surveying. In W. J. Lonner & J. W. Berry (Eds.), *Field methods in cross-cultural research* (pp. 85–110). Beverly Hills: Sage.

Mishler, E. G. (1986). *Research interviewing: Context and narrative*. Cambridge, MA: Harvard University Press.

Oakley, A. (1986). Interviewing women: A contradiction of terms? In A. Oakley (Ed.), *Telling the truth about Jerusalem* (pp. 231–251). Oxford: Basil Blackwell.

Patton, M. Q. (1990). *Qualitative evaluation and research methods* (2nd ed.). Newbury Park: Sage.

Seidman, I. E. (1991). *Interviewing as qualitative research: A guide for researchers in education and the social sciences*. New York: Columbia University.

SECTION IV

Future Reading

Chapter 12

Teachers and Research: An Annotated Bibliography

Dale T. Griffee

Seigakuin University

It is typical of an annotated bibliography to review a rather large number of books in an area and give one or two sentences describing each book. This article will review a rather small number of books because fewer books on research methodology are published than books in a content area but will review each book in more detail than is usual for most annotated bibliographies.

The purpose of each review is to give the reader a brief introduction to the book. Each review, therefore, can be considered only a snapshot of the book. While it is not possible to do justice to the complexity of any of these books anymore than it is possible to adequately evaluate a human being on the basis of a snapshot, it is hoped that an idea of the scope of each book plus some sense of the uniqueness of each book will assist the reader in making an informed choice when considering actual purchase. Toward that end, each review will strive toward description rather than prescription. The books are reviewed alphabetically by author.

A total of nine texts are reviewed. A book icon indicates the beginning of a review. The book title is given followed by the author or authors followed by the publisher and the year of publication then by the number of pages. The structure of each review is to list the citation of each book, then divide the review into four parts: the primary audience and the purpose of the book, the organization of the book, extra goodies, namely, features that are unusual or which might be overlooked, and the bottom line or a short summary.

This annotated bibliography consists entirely of books describing research methodology for teaching English as a second or foreign language. All the books are relatively easy to order and to purchase in Japan. All the books are TESL oriented, and no attempt was made to review books on research methodology from the mainstream content educational field of English speaking countries.

All of these books are helpful. For example they can help you understand research articles in the *TESOL Quarterly* and the *JALT Journal*. A classroom teacher who decides to engage in classroom research needs the experience of engaging in the research process, but in addition also needs the assistance of a guide or consultant. Without the experience of doing hands-on research, the new researcher has no grounding and the explanations in these books tend to sound abstract. Without the help of a guide, however, one tends to get lost. To put it another way, none of these books is a substitute for the experience of working with an experienced researcher. One of the most helpful experiences you can have is taking a course with the author of one of these books at a university or a TESOL summer institute. Furthermore, none of these books reads like a novel (i.e. most are not meant to be read from cover to cover). It is a good idea to approach these books with an actual problem or at least a question. It is also a good idea to read the relevant sections in more than one book as different authors explain their solutions in different ways.

Finally, keep in mind that there is no perfect book on research methodology. Each book reviewed here is a reflection of the author's context, purpose, and understanding of what research is and what new researchers need to know. All of these books are helpful either in whole or in part and there is no book reviewed here that would not be helpful to have in a teacher's personal library.

***Focus on the Language Classroom: An Introduction to Classroom Research for Language Teachers*.** D. Allwright and K. Bailey. Cambridge: Cambridge University Press, 1991. 250 pp.

Primary audience. This book was written in non-technical language for language teachers with little or no research background. The purpose of the book is to introduce classroom research to inexperienced teachers as well as experienced teachers who are not yet researchers so they can use research techniques to begin to think like researchers in the exploration of their own teaching. This is not, however, a "how-to book." Although including both, it stresses quantitative research over qualitative research, and provides a rather thorough review of research issues.

Organization. The book is divided into five main sections with two chapters for each section. Part one discusses the history of classroom research, part two

discusses what to take into account when starting research, part three discusses oral error, part four is about classroom discourse, and part five discusses student openness and what promotes or inhibits students' openness to learning.

Extra goodies. One of the surprises of this book, to me at least, is that it has two chapters devoted to error analysis. The authors seem to sense, probably rightly, that classroom language teachers are very interested in oral speech errors and how to research them. The background and the issues involved are thoughtfully discussed. The other goodie is the suggestions for further readings which follow each chapter. Each suggestion is a gem and reading them is like having an experienced teacher give you personal, cogent advice on what to read and where to look for it. These suggestions should not be overlooked by teachers new to classroom research.

The bottom line. This book is hefty, but readable. It is especially helpful for teachers in the process of deciding to take a research perspective in their classroom.

Understanding Research in Second Language Learning. **J.D. Brown. Cambridge: Cambridge University Press, 1988. 219 pp.**

The *primary audience* for this text is classroom practitioners, especially those who are terrified by numbers, charts, tables, and statistics. The aim of the book is to explain basic statistical terms; how charts, tables, and graphs work; and the appropriate use of research design. It is a book for the consumer rather than the producer of statistical studies.

Organization. There are 13 chapters which I, not the author, divide into four sections. The first section, chapter one, sets the tone for the book which is that while statistical studies are not the only kind of research or even the best kind, they are one type and they are useful. The second section, chapters two to four discusses variables and scales (nominal, etc.) which are used to talk about how research is basically designed. Knowledge of these scales will be useful later in the book because they are key to deciding what kind of statistics to use. The third section, chapters five to eight, gives a short background course in statistics. If you have had a good testing course, you can either skip these chapters or use them for review. The fourth section, chapters nine to thirteen, are the heart of the book.

Chapter nine, "Statistical Logic," explains the logic of statistical design. The next three chapters discuss the three major types of statistics: correlation, comparing means, and comparing frequencies. Each chapter has a section on choosing the correct statistic and a chart which uses variables and scales as a way of locating the statistic you might want to use. For example, if your dependent variable is a single interval scale and your independent scale is also a single

interval scale, where those intersect on the chart you find "simple regression" and "Pearson r," which are two statistical procedures you can use. All these terms and procedures are explained in the chapters in some detail.

Extra goodies. The special goodie for the new researcher is tucked away in chapter five which has the title "Critiquing Statistical Studies" because that is, in fact, the content of the chapter. What that title does not tell you is that there is a complete layout and description of the research article format. Research articles have a set format which cannot lightly be ignored. As a researcher you are expected to know the differences between such categories as materials, procedures, analyses, and results. You are also expected to know where an appendix is placed. Clearly laid out with each category explained and illustrated, this chapter alone is worth the price of the book. Check the format against the Publication Manual of the American Psychological Association, listed below, because there are slight differences.

The bottom line. A relatively easy to understand book for an introduction to statistics and the logic behind statistical studies. This is a book to refer to when you are at the research design level trying to decide which statistical procedure to use.

Research in the Language Classroom. **C. Brumfit and R. Mitchell. London: Modern English Publications in association with The British Council, 1991. 172 pp.**

Primary audience. This is not a textbook, in fact, it is not even a book. It is a collection of articles and the target audience is never explicitly stated. The purpose of the book, which is explicitly stated, is to survey major issues in classroom research. The influence of the British Council guarantees that this collection represents a British point of view. What saves this collection from being hopelessly parochial is that the contributors are widely distributed—Canada, Australia, Japan, U. K., and the U. S.; cover a wide spectrum of positions; and are highly qualified.

Organization. There are 12 articles divided into three major sections, references follow each article and there is no index. To give you some idea of the authors and articles, I will list the first five titles with their authors. These five articles constitute the first of the three major sections, General Issues: "The Language Classroom as a Focus for Research" by Brumfit and Mitchell; "The Teacher as Researcher" by David Nunan; "Ethnography: Bandaid, Bandwagon, or Contraband?" by Leo Van Lier; "Researching Classroom Language Learning" by Rod Ellis; and "The I-language Approach and Classroom Observation" by Vivian Cook. These five articles alone would justify the price of the collection.

Extra goodies. This is a rich collection, and I will comment on only four of the articles. The article by Brumfit and Mitchell is as fine an introduction to the

current issues in research as I have read. "Investigating Learners' Language" by Green and Hecht discusses in a very practical way some of the problems researchers will probably encounter. "Classrooms and Display Questions" by Banbrook and Skehan discusses theory-then-research (top-down) and research-then-theory (bottom-up) perspectives, and concludes with some suggestions on how the two could be combined. "Second Language Classroom Research and Teacher Education" by Michael Long continues a discussion introduced by Banbrook and Skehan by outlining problems of second language classroom research and, by implication, what we should be doing.

The bottom line. This is a valuable collection of articles because taken as a whole it deals with issues which are on the research agenda and which novice researches need to be aware of. Those issues are the teacher as researcher, types of research and their applicability, and a rationale for both quantitative and qualitative research design. How these issues are put and the answers that are given will form the new "common sense" of research and the research environment in which we will all have to live.

The Research Manual, Design and Statistics or Applied linguistics. E. Hatch and A. Lazaraton. New York: Newbury House, 1991. 628 pp.

Primary audience. The authors state that their text is for novice researchers, but that is like saying that the complete *Oxford English Dictionary* is for beginning readers. This book on statistics, research, and design is impressive. The purpose of the text is twofold: on the one hand, the authors want to promote statistical literacy as well as statistical production, and on the other hand, they want to promote research planning.

Organization. The book has 14 chapters and five appendices divided into five sections. Section one is on planning your research project, section two is on describing data, section three is on comparing groups, section four is on describing relationships, and section five (the conclusion) consists of five appendices which include statistical tables and lists of the formulas used in the book. Each of the 16 chapters has workbook features at the end of the chapter which includes activities and practice sections. If you are really a novice researcher it is doubtful you can understand these exercises without the help of a more experienced guide. On the other hand, it is not necessary to know the meaning of every word in the dictionary in order to read, even to read rather complicated material.

Extra goodies. In a book such as this, the category of extra goodies does not exist because the book is so vast and complex that most everything is not only there, it is systematically arranged. If, however, we change the topic of extra

goodies to I'm-glad-to-see-it, then I would nominate chapter 18 titled, "Assumptions of Statistical Tests." This chapter begins with a helpful discussion of reliability and validity, and concludes with assumptions of specific statistical tests as well as sections on solutions when assumptions cannot be met.

The bottom line. This is the mother of all ESL statistical research textbooks and if you are interested in a comprehensive discussion of the statistical procedures used in applied linguistics, this text should be in your library.

Approaches to Research in Second Language Learning. **D. M. Johnson. White Plains, New York: Longman, 1992. 253 pp.**

Primary audience. The primary audience of this text is graduate students in SLA, FL, and bilingual education. This is a first book and no previous knowledge of research methods or statistics is necessary. This book was written to provide an overview of the major approaches to research in second language learning. The author seeks to provide a balance between the quantitative and qualitative approaches.

Organization. Six approaches to research are discussed. Three approaches, correlational, experimental, and survey research are more quantitative while the other three, case studies, ethnographic, and multisite and multimethod studies are more qualitative. Two of the approaches will be discussed here.

Chapter 3, on "The Correlational Approach," is 22 pages long and contains an introduction, 9 criteria for analyzing correlational studies, and two sample studies both analyzed using the 9 criteria. Included in the introduction is a definition and example of a correlational coefficient and a discussion of validity and reliability in measurement, but no formulas or statistics are given. This is true for all chapters in this book. The one page conclusion discusses the major advantages and disadvantages of correlational research. The chapter ends with 30 reference citations.

Chapter 6, on "Ethnographic Research," is 27 pages long and is organized in a fashion similar to the chapter on correlation. The introduction discusses types and uses of ethnographic research, a section titled "Doing Ethnographic Research" discusses the role of theory, field techniques, and analysis. Ten criteria for assessing ethnographic reports are given with a sample study analyzed in detail using the criteria. The chapter ends with 65 reference citations.

Extra goodies. This book has many surprises. Each chapter has outstanding references. There is a chapter on the teacher-researcher movement and an appendix listing resources. The purpose of the chapter on teachers as researchers is to explore questions rather than propose answers. Questions are raised such as: Why do teachers see present research as irrelevant?, What is a teacher-

researcher?, and What are the goals of teacher-researchers? Do not skip the appendix, which is in three parts. First, it lists indexes and abstracts through which you can locate articles by topic or author. Then it lists reviews of research literature on particular topics, and finally a list of 33 journals each of which is briefly described.

The bottom line is that this is a good "getting started" book, especially if you are not clear on the various types of research. Because it covers the main types of research in EFL today, it might be a book to look at when you are at the stage of forming your research question.

Understanding Language Classrooms: A Guide for Teacher-Initiated Action. **D. Nunan. London: Prentice Hall, 1989. 183 pp.**

Primary audience. The target audience is teachers who are new to research. The purpose of the text is (a) to introduce educational—as opposed to linguistic—issues to ESL audiences, (b) to introduce the ins and outs of looking at classrooms, and (c) to present classroom research as a teacher renewal tool. Nunan uses (b) to accomplish (a) and (c).

Organization. The book is divided into seven chapters followed by three appendixes. After chapter one which gives a context, the remaining chapters group themselves into pairs. Chapters two and three are on who to look at (chapter two on teachers and chapter three on learners), chapters four and five are on methods of investigation or ways of looking (data collection and classroom observation methods), and chapters six and seven are on professional development or why look at all. Chapters three and four are especially important. Chapter three presents a readable discussion of three of the most important issues for teachers: SLA, tasks, and learner strategies, while chapter four discusses qualitative data collection methods such as diaries, field notes, interviews, and protocol analysis.

Extra goodies. Appendix C is an introduction to statistics. Statistics bother language teachers. If we had wanted to study mathematics, we would have become engineers. This appendix explains basic ideas as well as some current statistical application such as t-tests, analysis of variance, chi-square, and factor analysis. If these are strange sounding words, this appendix provides an introduction.

The bottom line. This is a basic book, but the seeds are all there. If the whole idea of research makes you ill, start with this book. If this book were taken seriously by a substantial number of language teachers, it would be a radical book with revolutionary consequences.

Research Methods in Language Learning. **D. Nunan. Cambridge: Cambridge University Press, 1992. 249 pp.**

Primary audience. This book is for practitioners in the field as well as graduate students. No previous knowledge of research is presupposed. The text was written to introduce research methods and tools, to prepare classroom teachers to understand and to critique research papers, and finally to help new researchers develop their own research strategies.

Organization. The book can be divided into two parts. The first part, chapters one to four, discusses research methods (experimental, ethnographic, and case study). The second part, chapters five to eight, discusses research techniques (classroom observation, introspective methods, elicitation techniques, and discourse analysis).

Nunan covers a broad range of issues and what you notice depends largely on what you are looking for. I appreciated his discussion of reliability and validity in chapters three and four, his discussion of surveys and use of questionnaires in chapter seven, and his discussion of discourse analysis, conversation analysis, and interaction analysis in chapter eight.

Chapter 10 concludes the text by discussing how to develop a research question and how to conduct a literature review—two problem areas for new researchers.

Extra goodies. This book has at least two extra goodies. The first bonus is a glossary of terms and the second is a chapter on program evaluation. Nunan provides a glossary of some 60 key terms used in research such as action research, ethnography, reliability, stimulated recall, and variance. This glossary is helpful to new researchers because these terms can not be found in a standard dictionary.

The second bonus is a chapter on program evaluation. Teachers have a strong stake in program evaluation, but generally they are not well trained in program evaluation which makes chapter nine welcome. Nunan's rationale for including this topic in a research textbook is that evaluation of an institution's program is, by his definition, research. Nunan discusses the difference between evaluation and assessment. He also includes a sample design for an evaluation study as well as an example of a case study.

The bottom line. This text summarizes information in lists and charts as well as giving clear examples and is especially helpful in its discussion of validity. This is a well balanced text which introduces major research areas and research techniques.

Publication Manual of the American Psychological Association **(4th ed.). Washington: American Psychological Association, 1995. 368 pp.**

Primary audience. Authors who prepare manuscripts for graduate level classes or for publication. There are many reasons to have this manual in your library: (a) Both JALT and TESOL follow APA which means that all manuscripts submitted to them must follow the APA manual, (b) in addition, most major journals in our field follow APA or a modified version of APA, (c) the manual is very complete and not only gives detailed examples of how to reference almost anything you can think of, it also gives instructions for issues such as use of abbreviations, quotations, numbers, footnotes as well as how to make figures and tables.

Organization. Seven major sections covering everything you ever wanted to know about writing a paper.

Extra goodies. Not applicable.

The bottom line. Don't fight it, just order it! You won't be sorry. Order from your bookstore or write to: Order Department, American Psychological Association, P.O. Box 2710, Hyattsville, MD 20784 U.S.A.

Second Language Research Methods. **H. Seliger and E. Shohamy. Oxford: Oxford University Press, 1989. 270 pp.**

Primary audience. According to the authors the book is intended for teachers new to research, but my impression is that it is not a text for complete beginners. This would be a text for a class with an experienced teacher or for new researchers with access to more experienced researchers who would be willing to act as mentors.

Organization. There are 10 chapters with chapters one and two serving as the introduction, and chapter 10 serving as the conclusion. Chapters three and four are a context to research, chapter five discusses the research plan, chapters six and seven discuss research design and chapters eight and nine discuss data collection and analysis. If I were to give this book a subtitle, it would be "formulating a research plan," in that, the authors believe that preparation is key to research. The text includes a thorough discussion of the differences between quantitative and qualitative research although one gets the feeling that in their hearts the authors prefer qualitative research. References are given after each section which sometimes causes certain articles to be referenced more than once. For example, Michael Long's article titled "Inside the Black Box" is referenced three times, once on page 112, once on page 134, and again on page 199.

Extra goodies. This is a tightly organized and interwoven text so that little tends to be overlooked. Nevertheless, I found the appendix reference for further reading to be helpful.

The bottom line. Reading this book is like hacking your way through a dense jungle and finding the treasure. While this book would not be in my small bag of 10 books to take to a desert island, it would be in my larger bag of 30 books. This is a standard text in the field by two very competent researchers, and it is unusual to read an article on second language research that does not reference this book. Key works to describe this book would be solid, turgid, insightful, and required. It is helpful when you are considering a research perspective or looking for guidance for a specific problem.

ISBN: 4-9900370-2-6 ¥2500

Classroom Teachers and Classroom Research

Classroom Teachers and Classroom Research is the second collection of articles in the JALT Applied Materials series. The articles in this volume are devoted to teacher-initiated research with a special focus on the classroom, which we recognize as the crucible of teacher concern and creative energy. While many articles are published in our field reflecting the research of teachers, this is the first volume published by JALT that is solely intended to define the role of teacher-initiated research in foreign language education, and help teachers become researchers in their own contexts.

Written by practicing classroom teachers in Japan, *Classroom Teachers and Classroom Research* includes:

- Language Teaching and Research ..... David Nunan
- Where Are We Now? Trends, Teachers, and Classroom Research ..... Dale T. Griffee
- First Things First: Writing the Research Proposal ..... Martin White & Diane White
- Literature Reviews: Obtaining Perspective ..... Paul Riley
- Methods for the Research Challenged .... Robert Homan & Chris Poel
- Designing a Language Study ..... James Dean Brown
- Action Research: Something for Everyone ..... Greg Hadley
- Ethnography Research ..... Linda Donan
- Designing Surveys for Language Programs ..... James Dean Brown
- Bare Minimum Knowledge for Understanding Statistical Research Studies. ..... Hiroto Nagata
- In-Depth Interviewing as Qualitive Investigation ..... Marilyn Books
- An Annotated Bibliography ..... Dale T. Griffee

Classroom Teachers and Classroom Research is edited by Dale T. Griffee and David Nunan, internationally recognized as an authority in classroom research and classroom research methods.

A Special Supplement to The Language Teacher

BEST COPY AVAILABLE